Treball de Fi de Màster

Master's degree in Automatic Control and Robotics

Combining Motion Planning with Social Reward Sources for Collaborative Human-Robot Navigation Task Design

WRITTEN RECORD

Author: Marc Dalmasso Blanch
Director: Alberto Sanfeliu Cortés
Date: April 22th, 2020

ETSEIB

Escola Tècnica Superior
d'Enginyeria Industrial de Barcelona

UPC

Abstract

Across the human history, teamwork is one of the main pillars sustaining civilizations and technology development. In consequence, as the world embraces automatization, human-robot collaboration arises naturally as a cornerstone. This applies to a huge spectrum of tasks, most of them involving navigation. As a result, tackling pure collaborative navigation tasks can be a good first foothold for roboticists in this enterprise.
In this thesis, we define a useful framework for knowledge representation in human-robot collaborative navigation tasks and propose a first solution to the human-robot collaborative search task. After validating the model, two derived projects tackling its main weakness are introduced: the compilation of a human search dataset and the implementation of a multi-agent planner for human-robot navigation.

A aquells en primera linia, sigui per elecció o necessitat.
Espero tenir un dia el valor d'acompanyar-vos.

CONTENTS

1. Introduction

Humans are social creatures. Human collaboration emerges naturally when interacting, not only to fulfil a task by exploiting the individual strengths of the participants or the group capabilities, but also as a mean to fulfil social needs. We collaborate on a daily basis and have multiple communication channels to support us, ranging from unconscious non-verbal signals to the most complex language registers. Likewise, we continuously infer each other goals and usually adapt our way to ease their achievement, obtaining the interaction and gratitude as the only reward. We are used to this behaviours to the point of forgetting their complexity and they shape our expectations when interacting with believed to be intelligent agents.

Human-robot collaboration is interesting from the complementary view. This becomes evident when focusing in manufacturing environments, where cobots provide precision, strength, specialized knowledge and indefatigability, whereas humans contribute with their versatility, adaptability, dexterity and common sense. In other settings, however, role assignment may be fuzzier or efficiency may be achieved due to group properties, depending on synchronous or simultaneous actions. Here, robots' communication and interacting capabilities, as well as their knowledge on human behaviour, may play a stronger role than their innate properties and skills.

In this thesis, we approach the human-robot collaborative search. We study the problem of searching for an object in a known simplified environment and approach this as a navigation-based task (overlooking physical interaction with the environment for occlusion removal). In this context, coordination means to explore the search space in a complementary fashion (i.e., avoid exploring the same or concurrent areas) to achieve the collaborative goal of finding the object. In this setting, humans keep track of other participants' behaviour and the whole task progress and infer other's future actions and goals. They plan exploration strategies taking into account all previous information, their influence on others actions and social norms, they even have multiple complex communication channels to support any need for explicit communication.

Collaborative solutions to such type of activities are most probably multi-modal. Should the robot follow the best solution and expect the human to adapt? Should it, on the contrary, enhance human comfort even at the cost of increasing task completion time? A plan may only be good if the humans act in a spectrum of expected behaviours, what if the human doesn't? How do we evaluate the goodness of such a plan? Should the robot and the human communicate? Through which channel? How do we represent the task knowledge for it to be understandable by both humans and machines? Conversely, should both of them focus on intention inference? How do we make the robot movement understandable? Answers to these questions are not trivial and, ideally, should most probably be personalised for each different human-robot pair or group.

In this work, we present a knowledge representation framework for human-robot collaboration. Over it, we develop a sampling-based motion planning approach to obtain feasible robot plans for human-robot collaborative navigation tasks. We test and validate both the knowledge representation and the planner interacting with humans in a virtual world and the system exhibits good collaborative performance. However, it strongly relies on human adaptive capabilities. To address this problem, we designed a Monte-Carlo Tree Search multi-agent planner using the same world representation. The current world health crisis, however, delayed the experimental campaign, so only a qualitative evaluation of the results is provided.

1.1 Motivation

On its strife for enhancing life quality, humanity has developed an uncountable number of technologies since the invention of the wheel. Through the years, we minimized the effort behind foraging, cultivation, harvest and manufacturing, yearning for comfort, available choices and, ultimately, free time. Automation of production and artificial intelligence might be the last technological stronghold to conquer to liberate humanity of obliged labor, relegating it to supervisory roles. Whether this turns out as utopia or dystopia remains a political struggle, but developing the tools is worth the shoot.

Human-robot collaboration points to be a core element in this scenario. Despite they may easily outrun humans on some applications, humans still remain as the core experts on many others. A great segment of population, however, lacks the knowledge to use such tools. May it be that they have a different field of expertise, are dependant or of early age or, simply, are not interested in them to such an extent. Robots will serve as tools for multiple purposes, hence they should be adapted to interact with humans as comfortably as possible for the latter. Making robots good collaborators will multiply their uses as well as empower the society.

As humans naturally do in their everyday life, in the future robots will regularly engage multiple agent navigation tasks. This ranges from efficiently organizing complex work labour tasks to reacting to spontaneous arising situations, as holding a door or deciding whether to pick up someone's wallet or leaving it to someone else who noticed it. In either case, stimulus or objectives create needs addressable by many agents, each one with different facility, required effort and motivation to undertake it. The ability to dynamically judge requirement and worthiness of taking action and judge others' intentions, as well as engaging dialogue when needed, will be a pillar of forthcoming social robots.

Mastering these abilities for collaborative navigation tasks is the first step towards this ideal. This covers from simpler tasks like map exploration, search or delivery to complex navigation logistics, as in resources management in construction projects or rescue operations. Human-robot collaborative navigation founds the stepping stone for all movement dependent collaborative tasks, thereby this thesis will focus on developing human-robot collaborative navigation models to advance towards this future.

1.2 Objectives

The main goal of this work is designing and implementing a model for knowledge representation, planning, and communication in human-robot collaborative navigation tasks. As many task designs and preferences may lead to the same task achievement, we propose a flexible multi-modal architecture. It should merge intuitive task design, general task-independent spatial communication, task-related contextual communication and theory of mind (knowledge models about the beliefs and intentions of the other agents). To sum up, the specific objectives of this thesis are the following:

1. Design and implement a complete spatial knowledge representation, able to express navigation tasks. We will try to look for intuitive representations comprehensible to non-expert users.

2. Design a planning algorithm capable of exploring the previous task representation, capable of achieving a fluent global collaborative task plan.

3. Maintain and update knowledge models for each actor involved in the task using the previous representation.

This models will be validated for a specific human-robot navigation task: the human-robot collaborative search of an object in an arbitrary space.

2. State of the Art

In this section, an overview of the related current state of the art is presented. Roughly speaking, we review human and robot task and goal representations, current approaches to human-robot collaborative navigation settings and several multi-agent and multi-robot works where we may find parallelisms to the human-robot collaboration case.

2.1 Task Representation

Human perception of the world is filled with both physical and abstract concepts. When tackling tasks, we construct and use abstract concepts such as roles, spatial relations, shifted importance of assets or the task itself. The same happens when a robot is designed to do a task, one programmer may define the tools or movements required or the task goals themselves, for example in the form of rewards (as they could also be specified in a descriptive manner, as formal logic propositions (state descriptors) to hold true, instead of a function or reward to optimize). However, such representations rarely fit. The subjective and qualitative perception of humans strongly differs from the objective and quantitative perception of the robot. The disparity might prove beneficial for collaboration. Nevertheless, when working together both should be able to convey and receive information about their environment, task or goals. Here, we review the current knowledge over human and robot abstract representations, as well as current approaches to represent and convey such knowledge in human-agent or human-robot settings.

2.1.1 Human Task Representation

We review human collaboration fundamentals and human-robot interaction challenges as we see them as the two founding pillars of human-robot collaboration. As this thesis focuses on human-robot collaborative navigation, we point our attention on pedestrian behaviour modelling publications and navigation approaches.

In [53] it is claimed: “perceiving and action planning are functionally equivalent, inasmuch as they are merely alternative ways of doing the same thing: internally representing external events”. Their theory of event coding (TEC) introduces a common coding system merging two classic action theory conceptual frameworks: the sensorimotor and the ideomotor view [44]. They treat all actions as goal-linked and equally in a common framework, including reactions and perception, which is defined as an active process.

2.1.1.1 Human Goals Representation

In 2002, Elliot & Thrash [30] examined “the role of approach and avoidance motivation in models of personality”. They compared spreadly used personality dimension models (i.e. Big Five and Big three models and the Eysenck’s traits) and affective disposition models (i.e. Tellegen modal and Watson and Clark model) with the binary model of approach and avoidance motivation. They observed that such two-factor structure remained robust for a variety of response biases and thus it is not simply a measurement-based artefact. They concluded that “it is reasonable to use measures of these basic dimensions as manifestations of or proxies for their corresponding temperaments”. They also insist stating that “much can be gained from interpreting the various literatures that have developed around each basic dimension through the lens of approach and avoidance temperament”. Their work supports the simplification of representing human goals as approach and avoidance motives, a concept frequently used (i.e. the social force model [49]). This representation is promising in the sense that multiple robot applications are build using the concept of cost or reward, quantitative scales of negative and positive feedback.

Goals are the main building blocs of human task representation. Their combination and dynamics, however, can be quite complex. Neal et al. [86] studies humans' dynamic self-regulation and the main properties of multiple-goal pursuit settings. In this work they explore human self-regulation through the adjustment of goal difficulty and the changes of goal importance over time, observing how they interact with processes that control the direction and duration of the effort (a result of the direction, duration and intensity of a task). They observe different factors influencing human leveraging of multiple goals (i.e. expected difficulty, incentives, environmental uncertainty, goal type -approach or avoidance- and valence) and explore different properties of multitasking settings (managing interleaved tasks). They found task switching behaviour to be common and to adversely affect performance, observe multifinality actions (enabling progress towards multiple goals) to be preferred over unifinal ones unless one of the goals is “highly activated” and discuss the human perception shift over the existence of goal shielding (tendency of an activated goal to inhibit accessibility to other goals). All these concepts may be taken into account when planning for human agents or inferring their objectives.

2.1.1.2 Shared Task Representation

Focusing on joint action [101], Bratman defined three characteristic features of any shared cooperative activity: mutual responsiveness, commitment to the joint activity and commitment to mutual support [13]. These properties, however, may take a gradual form. Sharing a conceptual common ground has huge implications in collaborative tasks, in particular when handling demonstrative references [107, 21].

One step ahead of having a common ground is building up a shared task representation. According to [111], shared intentionality transforms: “gaze following into joint attention, social manipulation into cooperative communication, group activity into collaboration, and social learning into instructed learning”. Human groups fostering the development of shared task representations are proven to outperform those who don't [116, 115]. Some research supports social agents are more prone to integrate other-generated actions in their task representations if they share a positive relationship [52] or there exists inter-dependency [98], although such conclusions are built upon the Simon Effect setup [104] whose social properties are currently challenged [27, 28]. As evidenced by research [4], people form shared task representations only when they perceive their co-actors as intentionally controlling their actions. Conversely, humans may form shared representations of tasks quasi-automatically when with perceived co-actors, even when it is more effective to ignore one another [102, 5, 113].

Implications of sense of agency [39, 83] and commitment [81] in task performance, uncertainty reduction and human experience, both in joint action [89, 99] and technology interacting situations [72], are currently being explored.

2.1.1.3 Human Social Behaviour

Several works focused on identifiable simple human social behaviours. It has been observed that humans imitate movements through observation-execution [32, 38], learn eye gaze connections to actions [16], use gaze dynamics like gaze-following or gaze aversion to express intention [56] and influence on the perception of others [90] and arm movement was used to predict human's hand goal [93]. Using a broader view, [74] detected and hierarchically clustered human actions, including facial expressions.

The study of pedestrian behaviour has a long history. In 1966, [47, 46] introduced the concept of proxemics, cultural dependent interpersonal distance relations. Later, [85] observed that only one-third of pedestrians do walk alone and explored the spatial organization of pedestrian groups and their impact on the overall crowd dynamics. Similarly, [23] focuses on pedestrian groups shape and their internal interpersonal distance. In this context, [80] integrated proxemics based physical and psychophysical features to train Hidden Markov Models (HMMs) to recognize spatiotemporal behaviours that signify transitions into and out of social interaction. A recent review on proxemics extensions, such as IPS [66], O-Space and P-Space, group formations and activity and affordance spaces can be found in [95].

Alternatively, some studies tried to directly model pedestrian dynamic behaviour. In 1995 Helbing and Molnar introduced the Social Force Model (SFM) [49, 48]. Many recent social navigation applications are based on this model or one of its extensions. For instance, [58] used evolutionary optimization to determine optimal parameter specifications for the social force model and [124] extended it by explicitly predicting the place and time of the next collision. These models rely on Gaussian decays, yet in [62] the interaction energy between pedestrians was measured, which suggested that interpersonal spatial relations followed power-law interaction based on their projected time to a potential future collision.

In [77], dynamic cost maps were learned through unsupervised learning to predict human movement. Not long after, [34] proposed a Bayesian human motion intentionality predictor able to predict pedestrian goals from a set of possible destinations given their dynamics. Alternatively, [114] proposed a framework to model the decision process behind human interaction-aware behaviour using non-cooperative game theory and the Nash equilibrium. Again focusing on pedestrians' interrelations, [119] presented an attention-based trajectory prediction model. In this perspective, [45] uses a simplified version of a multi-head attention mechanism to build a spatiotemporal graph that operates on the local and global contexts around pedestrians. Recently, human motion trajectory prediction is receiving a lot of attention, especially concerning machine learning approaches. A more extended review can be found in [97].

2.1.2 Robot Task Representation

Roboticists have used a variate range of methods to represent tasks. Ranging from task-focused implementations defining certain behaviours without complex knowledge abstractions to the building of semantic maps and extensive ontologies to represent the surrounding physical world and most abstract concepts. On the other hand, some fields defined a motivational representation of the world, assigning rewards to actions or events. A typical example of the latter would be all the family of reinforcement learning methods.

2.1.2.1 Task World Ontologies

A Task World Ontology is the representation of all objects physically present in the task environment or mentally present in the agent's mind during the execution of a specific task. In [117] they state that "The information should be recorded in a task model that captures relevant aspects of the users and their task world". According to [91] a planning task is a mapping from an eleven dimensional space to a plan model: initial world state, goal, plan task (parameters, preconditions, postconditions, achieved-by actions, required agents and time window), actions, agents, parameters, time horizon, constraints, preferences, cost function, solution criterion (complete and valid). Over it, any plan optimization should maintain pre-&-postconditions, condition consistency among task/actions and complete exclusion (two tasks can't occur at the same time if they are consuming the same agent). This definition, however, does not contemplate multifinality actions.

A recent research effort is being put to achieve a robot task ontology standard [7]. They work upon the assumption that actors need sufficient knowledge to perform actions, communicate activities and recognize and correct errors. One of their main objectives is the extension of the task world ontology standard to include collaborative task-specific concepts, as well as extend the complexity of constraint representation, such as functional requirements (e.g. "stay inside the safe zone"), temporal requisites ("complete the task in 10 seconds") or intertask dependencies.

2.1.2.2 Cost & Rewards

Despite the high-level knowledge representation efforts of previous works, the vast majority of real-world robot applications relay in more practical task representations. One may usually find task-focused repetitive motions or grid-like world representations for motion planning. One widely extended approach is to use the concept of cost or reward to establish a quantitative comparison between objectives and constraints.

Balance in cost or reward goal definitions usually is handcrafted ad-hoc or learned in a linear weighted model. Other approaches such as reinforcement learning use a predefined reward of certain world states or configurations and learn a non-linear reward distribution over the explored world states and action pairs. Many works achieved huge performance leaps with the latter approach, but the resulting model is usually highly dependant on the current world distribution and of uncertain behaviour on unexplored world states. This is extremely important when dealing with open worlds or extreme sources of uncertainty such as humans. As a consequence, some works explored the possibility of decomposing reward distributions, aiming to model-independent sources of functions of reward.

Lin et al. [73] explores the decomposition of rewards into sub-rewards obtained from different channels. They define that "reward decomposition views the total reward as the sum of sub-rewards that are usually disentangled and can be obtained independently" and observe how "the sub-rewards may further be leveraged to learn better policies". In their work, they maximize disentanglement between the multiple sources of rewards obtained from the decomposition, aiming to obtain its latent decomposition. Likewise, Štolba et al. [108] explores cost partitioning based on potential heuristics for multi-agent planning settings. They aim to provide a general technique for additive heuristic computation in multi-agent planning.

2.1.3 Task Representation in Human-Robot Teams

When facing human-robot joint action, it is of utmost interest to analyze disciplines as human-human joint action and connect them to the human-robot joint action case [51, 22]. In fact, according to [122] humans are capable of representing robot actions in a similar manner as they do to human's, in terms of action goals and means to achieve them. That, however, doesn't mean this representation is understandable by robots or, equivalently, that humans may understand the robots' internal knowledge representation.

Several works approached human-robot communication: Liu et al. [75] reviews gesture recognition techniques applied to human-robot collaboration, [110, 12] survey recent approaches capable of learning natural language and [79] reviews verbal and non-verbal human-robot interactive communications. On the other hand, other explore human goal inference and intent detection: Admoni et al. [1] reviews applications aware of the social eye gaze, [3] explores eye-hand behaviours in human-robot shared manipulation and [76] reviews intent detection, arbitration and communication aspects of shared control for physical human-robot collaboration.

2.1.3.1 Mental Models

Theory of mind approaches take importance as we try to model the subjective knowledge of different agents participating in the task. Estimating and maintaining mental states of other agents can reduce the number of unnecessary information given to the human [26]. Alternatively, Nikolaidis et al. [87] presents a game-theoretic model of the human partial adaptation to the robot, it decides optimally between taking actions that reveal the robot's capabilities to the human and taking the "believed as" best action.

Jonker et al. [60] studies metrics for measuring sharedness of team mental models. They define three quantitative metrics -model subject overlap (SO_i), model agreement (A_i) and sharedness (θ: $A_i \geq \theta \forall i | SO_i = 1 \forall i$)- and identify four main model components -ontologies, world state model, agent models and organizational specification-. In Chakraborti et al. [18], the authors state that achieving explicability implies either to conform to human expectations (generate explicable plans) or explain plans through a model reconciliation process. They differentiate four types of plan explanation processes characterised for the achieved model reconciliation: model path explanations, plan path explanations, minimally complete explanations and minimally monotonic explanations. Sreedharan et al. [106] aims to include the model reconciliation process in the planning phase, through generating self-explaining plans. These include actions responsible for explaining the plan itself.

2.2 Human Robot Collaborative Navigation

Human-robot collaboration is a field studying systems where humans and robots work together to achieve shared goals, a broad field sustained by many pillars: knowledge representation, planning, communication, plan sharing, decision making, agreement and adaptation. Human-robot collaborative navigation (HRCN) focuses on shared navigation tasks, ranging from goal allocation over a number of agents to complex synchronous movements, like dances or acrobatics. In this thesis we tackle human-robot collaborative search, thus we review current HRCN approaches in the state of the art. They are the first steps into collaborative models, but they are task-focused and thus can't be extended to other applications. In this paper, we go one step further considering the problem of collaborative exploration making use of a robot and a person.

2.2.1 Human as Manager

One of the fields tackling human-robot teams addressing navigation tasks is search and rescue robotics (USAR). The navigation of USAR robots working in unstructured environments is complex and many solutions use teleoperation and work semi-autonomous. In the Sherpa project [78, 103] they refer to the human as a "busy genius", who acts as a teleoperating commander, and work to ease the human's coordination of the whole multi-robot team. Recently, in [54], they achieve better performance in multi-robot search applying semi-autonomous teleoperation. Nevertheless, the need for methodologies enhancing human-robot teams collaboration has been broadly accepted [67, 121]. The TRADR project [68] aims to tackle this issue, highlighting report generation studies [63] and work agreement handling and evaluation [82], even though the agreement generation process is yet to be included. Their works study high-level human-robot interactions and draw interesting conclusions over experimenting with actual field personnel, though their robots are still being teleoperated.

In another field, Johnson et al. [59] present a new design process for human-robot collaborative applications focusing on identifying joint activity interdependences: the coactive design. They divide the design process into 4 phases: identification (interdependence analysis), determination of OPD requirements (observability, predictability and directability), selection and implementation. Other alternative approaches include co-driving, as the collaborative teleoperation of a robot through dialogue [36] or the collaborative control of wheelchair [15].

2.2.2 Human-Aware Navigation

There is debate whether navigating through a social area (with pedestrians) is a collaborative task or not. Supporters of this idea identify core collaborative aspects in navigating through this environments: needing to infer others actions and goals, take into account the own influence on others, ease other's achievement to clear own's path... while detractors, on the other hand, argue that though this action implies many complex social interactions, it does not define a common goal, so that many core components of collaboration are missing. This discussion falls out of the scope of this thesis, but it's out of the question that human-aware planning did ground the basis for many human-robot collaborative navigation core pillars.

Lasota et al. [70] presents a human-aware motion planner that not only improves perceived satisfaction but also leads to more fluent teamwork, more concurrent motion and shorter human and robot idle times. [29] states that functional movement can harm coordination. They defend that legible motion, proactively conveying its goal, performs better in human-robot collaboration setups than predictable motion, expected known the goal. Later, in 2017, [9] presented Sociosense, a social navigation model using psychological profiling of pedestrians further extended in [8], where they use CNN-based learning and the PAD (Pleasure-Arousal-Dominance) model from psychology, classifying pedestrian characteristics into four emotion categories (happy, sad, angry, neutral) and applying emotion-based proxemic constraints. Meanwhile, several approaches for crowd navigation based in deep reinforcement learning have been published [31, 19, 20].

The reviewed models focus on general crowd human-aware navigation, but some approaches tackle specific frequent testbeds in social navigation. [123, 55] present methods for human following. [25] concludes that, generally, seated humans prefer to be approached from the sides, [100, 61] present strategies to proactively approach customers in a shopping mall and [92] uses inverse reinforcement learning to approach people. [69] uses directional cost models to enhance robot behaviour when in crossing situations and [65] optimizes a graph representation for trajectories pursuing the same objective. Assistive robotics writes its own chapter in human-aware navigation. [118] presents a human-aware model for wheelchair navigation and, later, [11, 84] present human-aware models taking into consideration both pedestrian and user comfort. Other works approach cooperative social robots to accompany guided groups of people [41] or search and track of people through believe modelling [42]. A more extended review in social navigation approaches can be found in [95, 64].

Side-by-side navigation is a special case of human-aware navigation. [33, 35, 40] approach this challenge through SFM-based methods and, in parallel, [109, 57] present methods for side-by-side wheelchair navigation. They are the first steps into collaborative models, but they can't be extended to other applications as they are task-focused. We pursue a flexible model capable of representing multiple tasks and conveying such representations to the human. Moreover, HRCN models should integrate human-robot communication, intentionality detection and inference, shared planning, reasoning, decision making, agreement and adaptation mechanisms.

2.2.3 Metrics for HRCN

We find no homogeneous metric definition in human-robot collaborative settings, though some trend begins to arise. Chakraborti et al. [17] discuss and compare several frequently used metrics for motion planning in human-robot shared environments. They define the difference between legible, explicable and predictable plans (defining the three as gradable spectrums), consider the difference between implicit and explicit explanatory actions and discuss the reasons behind plan and goal obfuscation. They maintain that a motion planning framework for such environments "should not only be able to compute plans but also policies for communicating its information content during execution".

Though not directly related to navigation tasks, there are other works considering metrics for human-robot collaboration. It has become necessary to quantitatively analyze the performance of the heterogeneous teams to enable comparison between different team configurations. [37] presented a formulation for a decision-analytical based measure of trust. Additionally, [105] made a survey on quantitative team performance metrics for HRC for space exploration missions. Recently, [50] reviewed present subjective and objective fluency metrics for physical human-robot collaboration (PHRC). He suggests to carefully observe objective metrics dynamic behaviour, given their variability, and studies their correlation with subjective metrics. To do so, he defines four objective metrics:

- **Human Idle Time (H-IDLE):** Percentage of the total task time that the human is not active. Observed to be significantly correlated with subjective fluency.
- **Robot Idle Time (R-IDLE):** Percentage of the total task time that the robot is not perceivably active. Found to be consistently inverse correlated with fluency, significance limited to very high values.
- **Functional Delay (F-DEL):** Accumulated time between the completion of one agent's action and the beginning of the other agent's action. Observed to be significantly reverse-correlated with subjective fluency.
- **Concurrent Activity Time (C-ACT):** Percentage of the total task time during which both agents have been active. On very low values leads to a drop in fluency perception.

Other commonly used metrics are: completed tasks, time, bandwidth used, time to complete, duplicated effort, total number interferences, number of exchanged messages ([59, 120]).

2.3 Multi-Agent Planning (MAP)

When tackling a collaborative task for human-robot teams, one may find inspiration in multi-agent or multi-robot implementations for the same or similar challenges. These approaches will lack many required components of HRC but can offer interesting insight into successful individual robot behaviour and team distribution strategies. An extensive survey reviewing MAP algorithms may be found in [112, 96]. In both one may find interesting taxonomies for MAP solver classification, where the former catalogues different approaches over these categories: agent distribution(planning/executing agents), computational process(centralised/distributed), plan synthesis scheme (unthreaded/interleaved), communication mechanisms, heuristic search (local/global) and privacy preservation.

2.3.1 Task Allocation

Works focusing on task allocation tackle the challenge of planning over a set of tasks presenting strong temporal and ordering constraints. They focus on the planning process, hence usually evaluate their results through the viability and optimality of their plans or test them on simplified world representations. Though such evaluation isn't viable in human-robot settings, it's interesting to study how do they tackle the complexity of planning over such restrictions. Nunes et al. [88] presents a taxonomy for task allocation problems with temporal and ordering hard and soft constraints, which they may identify as time window, synchronization or precedence constraints. Their review of commonly used optimisation objectives is of special interest: miniSUM, miniMAX, miniAVE (average), minimize lateness or tardiness, minimize idle time, maximize the number of tasks completed, minimize the number of robots used and maximize profit ($rewards - costs$, utility).

Some interesting recent works tackling this challenge for navigation tasks may focus on handling path conflicts [125] or generate plans through a multi-population genetic algorithm [6]. Jonker et al. [60] presents a high-level theory of mind planner for multi-agent search and retrieval working over the simplified world simulation BW4T. On the other hand, Wei et al. [120] works over a similar setting while testing the effect of different communication types over the overall task performance. They concluded that communication has a higher impact when there are temporal constraints between goals, communicating beliefs over the task world is generally more costly than communicating goals, but achieves greater performance, and communicating both beliefs and goals has no significant improvement on performance over only using the former.

2.3.2 Multi-Agent Path Planning

In classical path planning, sampling methods is a widely used family of algorithms to retrieve good paths. Their strong points are being anytime, capable of incorporating prior knowledge and suitable for online replanning. Rapid random trees (RRT) and probability road maps (PRM) are well-known examples of this family. In multi-agent planning, however, there is another family of sampling algorithms that grows in protagonism: the Monte Carlo Tree Search (MCTS). Browne et al. [14] reviews the usage of these methods and a large set of possible variations and enhancements present in the literature. One good recent example of a multi-robot MCTS implementation for path planning is the work of Best et al. [10]. They adapt the MCTS to an interleaved distributed setting while presenting a new tree expansion policy (discounted UCT). In this work, each robot expands a global tree for all the agents and shares its computed probabilities for a set of dynamically selected available plans over a PRM connected through Dubins paths. Their implementation enables a robust distributed global plan expansion. Likewise, Li et al. [71] also approaches distributed multi-agent planning through MCTS.

2.3.2.1 Collaborative Exploration

In the mobile robotics field, one fundamental problem is the exploration of an environment. The most important issue during exploration is where to move the robots to minimize the time needed to fully explore the environment. Whereas the exploration problem has been studied in detail for single robots [2, 43], there are only a few approaches for multi-robot systems [94].

3. World Knowledge Representation

In robotics, it is useful to model robot actions over the expected reward. This concept relates to human beings, as the expectancy of satisfaction or regret motivates our actions [30]. Humans identify and model the sources of such feedback (from now on reward). Our tastes and preferences shape our view of the world and we adapt our behaviour to maximise them [86]. The social reward sources (SRS) idea is to model the robots' world representation as a set of reward sources $\psi \in \Psi$ to build functional human mental model representations and ease human-robot knowledge sharing.

In open worlds, the addition or subtraction of a new element in an environment may not necessarily change the task at hand or change one actor interaction idiosyncrasy versus the other in the scene[49]. Likewise, its influence may be locally bounded [47]. Other works have suggested the existence of disentangled sub-rewards in learned behaviours [73], some proposing their combination in an additive fashion [108]. We believe a modular representation of the sources of reward present in the task world can provide a flexible task representation, capable of adapting to changes in the environment and the actors.

3.1 Social Reward Sources

Following the advices expressed in [53] and [122], we want to use SRSs to build an unified subjective world representation of all external events Ψ_W (Fig. 3.1). This includes perceived and interpreted environment influences E (from avoidance of objects like walls or furniture to conceptual abstractions such as areas in a sports field), the pleasantness or effort to perform certain actions A, tasks and goals we may engage T and interaction feedback emerging from the conjunction of all agent's actions I. All these concepts may be modelled as sources of reward (Ψ_E,Ψ_A,Ψ_T,Ψ_I) and seen from a different subjective perception by each agent (e.g. risk, capabilities, effort, personal preferences or eagerness to do a task or to follow someone's instructions). When task planning over the SRS model, one or various tasks with different goals may coexist in the same world representation.

$$\Psi_W = \{\Psi_E, \Psi_A, \Psi_T, \Psi_I\} \tag{3.1}$$

In multi-agent settings, one may consider multiple agents are engaged to the same task or affected by the same circumstances. This means each source of reward $\psi_n \in \Psi$ may influence multiple agents, from now on identified as the set of targets τ_n. Likewise, one may define the task world for

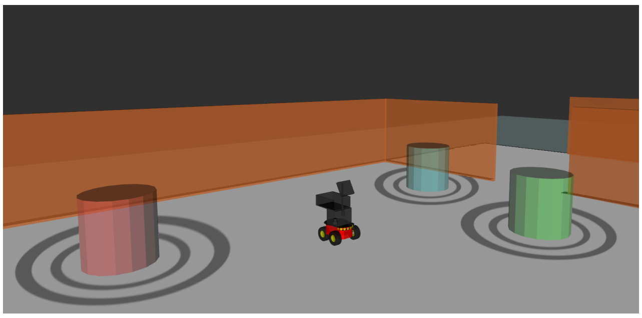


Figure 3.1: **SRS Model.** This model aims to achieve a common world representation for the physical world perception, task modelling and human-robot communication. A human understandable representation that eases knowledge sharing and permits direct navigation planning over it.

each agent, maintaining a mental model of their knowledge. Shared goals, multifinal actions and capabilities may be represented through the consumption, stacking and shaping of independent reward functions.

$$\Psi = \{\psi_0^{\tau_0}, \psi_1^{\tau_1} ... \psi_k^{\tau_k}\} \tag{3.2}$$

3.1.1 SRS Definition

Ultimately, a social reward source $\psi \in \Psi$ is a generative model that defines a reward function along all the search space $r(\psi) = f(x,y)$ (this work focuses in $\mathbb{R}^2$ navigation, but could be applied to $\mathbb{R}^3$ or a robot joint space, for example). Nevertheless, while humans may easily relate to concepts like attraction and repulsion, our world abstractions are complex and often related non-trivial abstract shapes and volumes (e.g. activity, affordance and information processing spaces from proxemics theory [95]).

Flexibility is a must if we are interested in being able to describe the full range of possible human dynamic, spatial and demonstrative communication. Humans understand symbolic level abstract concepts as "room" or "flat", as well as relative quantification of gradable language (e.g. rather, quite, very or dreadfully urgent) and relative positioning of objects. Both robots' task and planning level representations should be able to adapt to incorporate this information, including shape, localization and strength. Following this principle, some works use 3D object representations as communication channels. For example, Johnson et al. [59] builds the human-robot communication interface around 3D visualization and click-&-drag interactions of virtual objects, named manipulables. They observed that "manipulables proved so valuable that they where consistently used", their virtual arms representation ended up being used in 99% of all arm commands.

Consequently, we find it useful and more compact to take position and shape (in $\mathbb{R}^2$) as core characteristics of the SRSs. In $\mathbb{R}^2$, sources may be linked to an *object*, *area* or *point*, having their shape and boundaries parameterized (point sources are taken as punctual boundaries). In these cases, their *reward models* f are defined in the boundary reference as function of the distance to the nearest boundary point $r(\psi) = f(d_b)$ (e.g. constant values or Gaussian and exponential decays, Fig. 3.2). Moreover, object and area sources may present two different reward models for internal and external zones.

On the other hand, some rewarding or troubling situations may affect oneself continuously, such as proximity to hazards, while some may only affect during a time window, until some task completion or instantly due to some event (e.g. pressing a button, crossing a crossroad or holding a door to someone). We identify a special type of sources, which we name consumable sources. These are sources whose time dynamics are affected due to actors actions, usually being "consumed" over

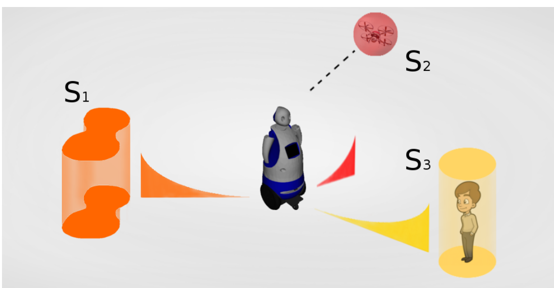


Figure 3.2: **SRS Properties.** Linking reward functions (positive or negative) to physical entities or spatial abstractions enables their modeling over meaningful relative distances.

time. These take special importance in the multi-agent case, where the modelling of such dynamics differentiates between global goals needed to be met by all target actors and goals achievable by one or more actors, while possibly undertaken by or meaningful to a higher set of targets.

Summarizing, each reward source is defined by the following characteristics: *type* (repulsive or attractive), *temporal behaviour* (constant, variable or consumable), *reward model* (e.g. gaussian or power function decays), *geometric properties* (e.g. object, area, point) and *targets*.

3.2 Communication using the SRS model

In human-robot collaboration, SRSs act as common knowledge representation building blocks. If both agents grasp the SRSs concept it is possible to achieve a shared task representation, understanding each other's intentions and task contribution. In these circumstances, it becomes easy to share knowledge using the same building blocks. Its modular definition and it being capable of handling multitasking allow a fluent integration of other's knowledge and instructions. Moreover, its decentralised design is tolerant to communication dropout.

Sharing knowledge and task representations also eases maintaining an estimation of others' knowledge and perception. This encourages the proposal of plans that take into account capabilities and preferences of each agent, plus the model provides compact information encoding. For example, in the proposed application the robot keeps a model of the believed perceived exploration by the human.

All these virtues, however, build over the assumption of the model being intuitive for humans. To test this property, we have performed a user study using the model ((Section 6.1). This study proves that non-expert humans can understand the main properties of this knowledge representation and use them to design robot behaviours.

4. Motion Planning in the SRS World

Once all relevant information is represented in the SRS model, we should find a suitable navigation plan along the reward space. There are many possible approaches to explore the SRS world representation. One may consider discretising the space and apply classic methods as the $A^{\star}$, but at the cost of losing the richness of the continuous rewards space representation. Conversely, many novel reinforcement learning approaches could be suitable for this reward space exploration, though it would be a challenge to extract the underlying knowledge and share it with the humans. Sampling-based algorithms, on the other hand, offer simplicity and any-time solutions while working over a continuous world representation.

There exists vast literature addressing the problem of planning and generating motion from a reward distributions. Some approaches focus on the present. Markov Decision Processes (MDPs) try to learn an optimal policy for action selection given while, possibly by learning an underlying value function of states or action-state pairs. Other approaches, however, try to extend the decision making towards the future. Classic sampling methods such as RRTs and PRMs or originally game theoretic approaches like MCTS have many similarities and may generate long paths (or decision trees) into the future. When trying to develop a motion planner, one should keep in mind its use objectives. A HRCN plan should have enough time depth as to reach meaningful shared goals and be shared in a comfortable pace. Although we decide to keep the time horizon variable, we expect to receive plans of a considerable length.

4.1 Regarding Goals

We aim to generate paths over a set of multiple reward functions, thus we expect them to be potentially multifinal (i.e. serving multiple goals or purposes). In motion planning, however, literature often commits to a single goal (the actual end point of the movement sequence, or *where*) while framing the rest of objectives as constraints or costs (path shaping directives, or *how*). The goal there may become a core characteristic of the planner, being selected on a higher task allocating layer, and focus on finding the optimal path from the homotopic set (e.g. splines and elastic bands).

We, however, want to leave the goal undefined, fuzzy, and select it indirectly through the most rewarding path. These may lead to less precise paths, due to the search space increase, but delivers a pragmatic approach to global planning in a multitask space while performing a space informed task allocation. (It is important to note that tasks' costs in sequential navigation tasks are directly related to their predecessors location and the environment layout, information that may not be available, specially in open world settings).

This brings us to a convention contradiction: assuming infinite planning time, should we let the planner expand until finding a path sequencing all goals? In a collaborative setting we expect the agent partners to contribute to the task progress. The task world will experience frequent unpredictable changes, so we are bound to constant replanning and such a plan looses significance. Thus, we decide to treat navigation goals as rewarding end-path locations, generating paths that focus on clear significant goals. We call them *path selection* attractive sources, as they have no substantial effect on the path generation phase (which rather aims at covering the search space). Nevertheless, we may have positive constraints over the path shape, zones through which paths are encouraged to go through. We call rewards encoding such constraints *path shaping* attractive sources. In sampled-based path planning implementations we differentiate between the path generation (or tree expansion) and the path selection phases and *path selection* sources may be encoded to affect only the latter. It is important to remark that it responds to a path planning convention and aims to simplify the planner outputs to ease human comprehension.

4.2 Motion Planner

For simplicity and computational efficiency, we choose to use an adaptation of the well-known $RRT^\star$. In our implementation, cost functions may be dynamically updated along each branch expansion of the tree.

4.2.1 SRS as a Collision Avoidance Soft Constraint

Conceiving obstacles in the environment as negative reward sources (or repulsors) builds up a particular costmap that shapes paths taking into account human perception of objects proximity (Fig. 4.1). Until now, such awareness has only been taken into account in local planning schemes. Adding it to global planning is an attempt to, first, establish a framework coherence and, second, avoid planning under unrealistic assumptions. Socially navigating through a hallway with certain distributed objects may be more time consuming than taking a longer path, while this might not be the case for standard navigation.

It is important to note that these SRS sources are assumed to affect the robot continuously. As a consequence, they are defined as a density and applied over the RRT nodes connection. Each node cost would be defined as follows:

$$C(P) = \sum_{i}^{P} \left(C_i + c_{i-1,i}\right) \tag{4.1}$$

$$C_i = \sum_{j}^{|\Psi|} (\psi_j(i) \cdot d_{i-1,i}) \tag{4.2}$$

where $C(P)$ is the cost contribution of Ψ to the path P, being Ψ the set of repulsive force sources in the environment. Notice that $\psi_j(i)$ stands for the reward given by source ψ_j to the path node $i \in P$. Finally, $c_{a,b}$ and $d_{a,b}$ denote respectively the movement action cost and the distance from a to b, being the latter in the dimensional magnitude over which constant cost densities are defined.

4.2.2 Path Shaping Tasks

The previous extension may generate socially acceptable paths, which is a useful passive property but turns out to be limited when considering the design of navigation tasks. If we attempt to create a framework for flexible human-friendly design of navigation tasks, it should be capable of generating

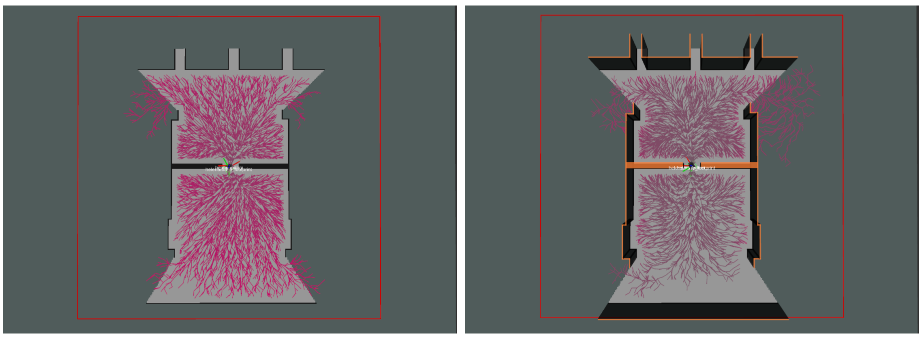


Figure 4.1: **SRS as a Collision Avoidance soft Constrain.** Euclidean cost versus environment objects SRS obstacle avoidance.

non-trivial path shapes related to the task at hand. One good example of this is the previously introduced concept: “in your way, pass through this zone”.

Now, to tackle this problem the concept of constant attractors was considered, a source of negative cost which is known to cause undesired behaviours as in Fig. 4.2.a. Creating a repulsor equivalent to the Gaussian attractor is neither a good idea, as it is equivalent to increase the movement action costs along the rest of the space, thus the effect of the other repulsors is hindered proportionally to the magnitude and number of attractors added (Fig. 4.2.b and Fig. 4.2.c). We assign to each path the maximum sampled value of the source. Note that this is equivalent to dynamically decrease (or consume) the source reward on each tree branch so asymptotic optimality is no longer ensured, a property already lost when considering open environments with dynamic objects.

The final RRT expansion cost is defined as follows:

$$C^{gen}(P) = C^{gen,ct}(P) + C^{gen,cs}(P) \tag{4.3}$$

$$C^{gen,ct}(P) = \sum_{i}^{P} \Big(\sum_{j}^{\Psi^{gen,ct}} (\psi_j(i) \cdot d_{i-1,i}) + c_{i-1,i} \Big) \tag{4.4}$$

$$C^{gen,cs}(P) = \sum_{j}^{\Psi^{gen,cs}} (\max_{i}^{P} \{\psi_j(i)\}) \tag{4.5}$$

given $C^{gen,k}$ is the cost contribution to the path P of $\Psi^{gen,k}$, the set of path generation sources ψ of nature k (possibly being constant ct or consumable cs).

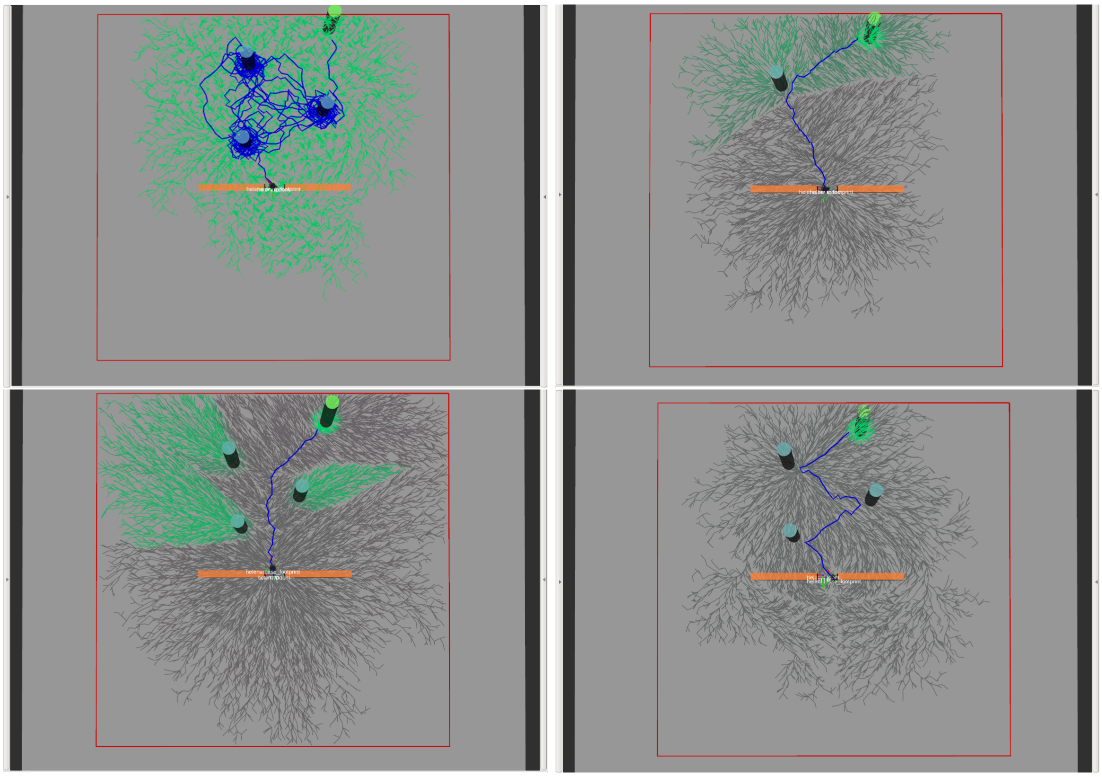


Figure 4.2: **Consumable Sources.** From left to right, up to down: a) Attractive cumulative cost behaviour, b) one equivalent repulsor hindering objects effect, c) multiple equivalent repulsors hindering each other’s effects and d) consumable source approach.

4.2.3 Path Selection

Cost formulation in path selection phase is equivalent to that of path generation but taking into account path selection sources. The general formulation may be written as:

$$C^{sel}(P) = C^{sel,ct}(P) + C^{sel,cs}(P) \tag{4.6}$$

$$C^{sel,ct}(P) = \sum_{i}^{P} \Big(\sum_{j}^{\Psi^{sel,ct}} (\psi_j(i) \cdot d_{i-1,i}) + c_{i-1,i} \Big) \tag{4.7}$$

$$C^{sel,cs}(P) = \sum_{j}^{\Psi^{sel,cs}} (\max_{i}^{P} \{\psi_j(i)\}) \tag{4.8}$$

Likewise, $C^{sel,k}$ is the cost contribution to the path P of $\Psi^{sel,k}$, the set of path generation sources ψ of nature k. This implementation allows an easy generation of paths like the one shown in Fig. 4.3

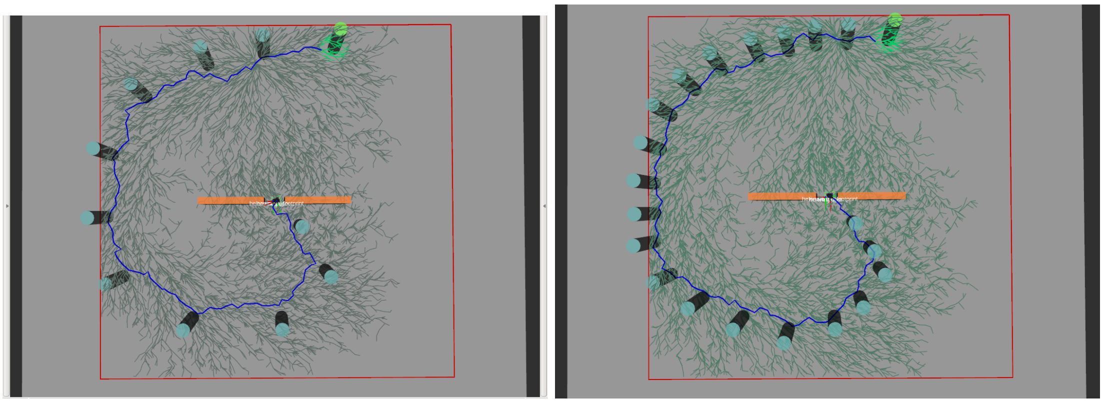


Figure 4.3: **Path generation trough SRS.** An example of a path generated through SRSs. It shows the robustness of the consumable source approach against changes in the number of sources.

5. Collaborative Search

One recurrent human activity is searching. Either due to lack of memory, unintended loss or a third party intervention, we usually find ourselves looking for something. For instance, it is frequent to see people looking for their keys before leaving their house in the morning, often on a rush. In that situation, it's not strange for them to ask for help or information while already engaging the task in parallel. Moreover, the strategies followed vary through time and are strongly influenced by their beliefs about the object location. In the keys example, pants' pockets, bedside tables and door locks may be the first locations to be checked.

Here, we approach the challenge of designing robots capable of collaboratively participate in search tasks while being part of human-robot mixed teams. In general, searching may include a large spectrum of actions such as active perception and object manipulation. In this work, however, we consider a pure navigation task where the environment is assumed to be fixed (no possible physical interaction with the environment for occlusion removal), the robot sensors are static in the robot platform reference and the target object is on ground level. On the other hand, both the robot and the human can move freely, both team members may follow any search policy and even lose contact.

5.1 Problem Statement

The human-robot collaborative search task may be defined as the process through which a given team of agents A, comprised of both humans and robots, explores a known space to locate an object O. During this process, agents can update their belief over the object location through the exploration of the environment, the observation of their colleagues' actions and the received information through active communication. The task is assumed as finished when the object is found.

In the experiments presented in this thesis, the team consists of one person and one robot. We focus on searching both human and robot accessible spaces, so both are assumed capable of navigating through it autonomously. To perform this task, we build an observability graph upon a discretised representation of the search space, as shown in Fig. 5.1.

In the following sections, we formalise the core mechanics of this problem.

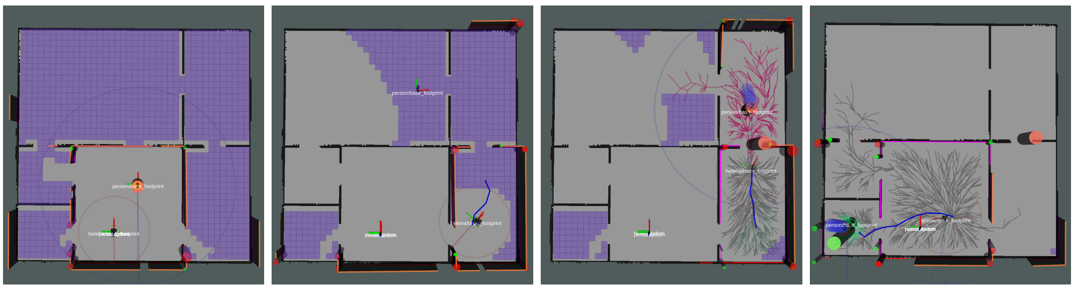


Figure 5.1: **Collaborative Search Testbed.** From top to bottom, left to right: a) The robot infers the unexplored zone from its detection range (red circle) and the person's (blue circle). b) People detection is impossible when the person is out of sight, hence no inference is done. c) The person indicates the robot to avoid searching through that zone, as either it is already explored or the person will do it on their own. d) The person finds the object, thus indicates the robot to come.

5.1.1 Agent Detection Model

During the search process, the object location belief is continuously updated based on the actors' actions. Hence, we should model the probability of an agent $a \in A$ detecting an object O at a certain location $\vec{p}$.

$$P(D_a(r_a, \theta_a, \Delta t)|O(r_a, \theta_a))$$

where $\Delta t = t_f - t_0$ is the search time, (r_a, θ_a) are the polar coordinates of the location p in the human a reference and $D_a(r_a, \theta_a, \Delta t)$ and $O(r_a, \theta_a)$ state the object being detected by agent a and actually being at the given location, respectively. We make a number of assumptions:

Assumption 1. Detection models are independent of their initial time t_{0_i}. In other words, human detection capability does not change over time. So for one agent:

$$P(\overline{D}_a(\vec{p}, \Delta t)|O(\vec{p})) = \prod_i P(\overline{D}_a(\vec{p}, \Delta t_i)|O(\vec{p}))$$

where the overline in $\overline{D}_a$ expresses the complementary statement, i.e. being "undetected by agent a", and

$$t_{0_{i_0}} = t_0, \quad t_{f_{i_f}} = t_f, \quad t_{f_i} = t_{0_{i+1}} \quad i = i_0, ..., i_{f-1}$$

Assumption 2. Detection models are independent of $\vec{p}_a \forall a \in A$. Human detection capability is independent on the perceiving human position, as long as (r_a, θ_a) is visible, and on all other participating agents' position. Whereas, change of focus or occlusions in its field of view due to other teammates proximity are not considered. So for each location:

$$P(\overline{D}(\vec{p}, \Delta t)|O(\vec{p})) = \prod_{a=1}^{A} P(\overline{D}_a(\vec{p}, \Delta t)|O(\vec{p}))$$

5.1.2 Object Location Probability

At a given time t, where t_0 is the task beginning and $t_f = t$, the updated object probability on each location given the current accumulated global search is:

$$P(O(\vec{p})|\overline{D}_t) = \frac{P(\overline{D}_t|O(\vec{p})) \cdot P(O(\vec{p}))}{P(\overline{D}_t)}$$

To update the global object location belief, we make the following assumption:

Assumption 3. Humans make no false positives while searching or, on another perspective, they filter them automatically.

$$P(\overline{D}_t|O(\vec{p})) = \prod^{\vec{q}} P(\overline{D}_t(\vec{q})|O(\vec{p})) = P(\overline{D}_t(\vec{p})|O(\vec{p}))$$

$$P(\overline{D}_t) = 1 - P(D_t) = 1 - \sum^{\vec{q}} P(D_t(\vec{q}))$$

$$= 1 - \sum^{\vec{q}} \sum^{\vec{p}} P(D_t(\vec{q})|O(\vec{p})) \cdot P(O(\vec{p}))$$

$$= 1 - \sum^{\vec{p}} P(D_t(\vec{p})|O(\vec{p})) \cdot P(O(\vec{p}))$$

consequently:

$$P(O(\vec{p})|\overline{D}_t) = \frac{P(\overline{D}_t(\vec{p})|O(\vec{p})) \cdot P(O(\vec{p}))}{1 - \sum^{\vec{p}} P(D_t(\vec{p})|O(\vec{p})) \cdot P(O(\vec{p}))}$$

The previous formula can be further simplified for the uniform prior case. If we are working on a uniform space discretisation, we may use:

$$P(O(\vec{p})|\overline{D}_t) = \frac{P(\overline{D}_t(\vec{p})|O(\vec{p}))}{N - \sum^{\vec{p}} P(\overline{D}_t(\vec{p})|O(\vec{p}))}$$

where $P(\overline{D}_t(\vec{p})|O(\vec{p}))$ is iteratively updated from observations using the detection model. One may obtain an efficient belief using a dynamic programming approach.

5.2 Task Modelling through SRS

The SRS modelling of the collaborative search task Ψ_{CS} is quite straightforward. We represented the collaborative search as a number of SRSs equivalent to that of the search space discretisation. In other words, each possible object location generates a function of reward. Each source ψ_L generates a reward proportional to the probability of detecting the object on the source location $P(O(\vec{p}_{\psi_L})|\overline{D}_t)$ along the search space (Fig. 5.2.a). All sources can be combined by the planner, for instance in an additive form, while being subject to independent consumable dynamics. This allows for a coherent reward evolution inference on the planner simulations whilst allowing the planner to be disentangled from any task knowledge, such as visualization restriction or detection models.

$$\Psi_{CS} \quad = \quad (\psi_L \quad | \quad P(O(\vec{p}_{\psi_L})) > 0) \tag{5.1}$$

Likewise, belief is updated along the search progress and sources are generated at each replanning phase. Fig. 5.2.b shows a timestamp on a two-people search process. On the other hand, Fig. 5.2 is the visualization of merging all the resulting social reward sources, which is proportional to the probability of seeing an object from the given lookout (discretisation block).

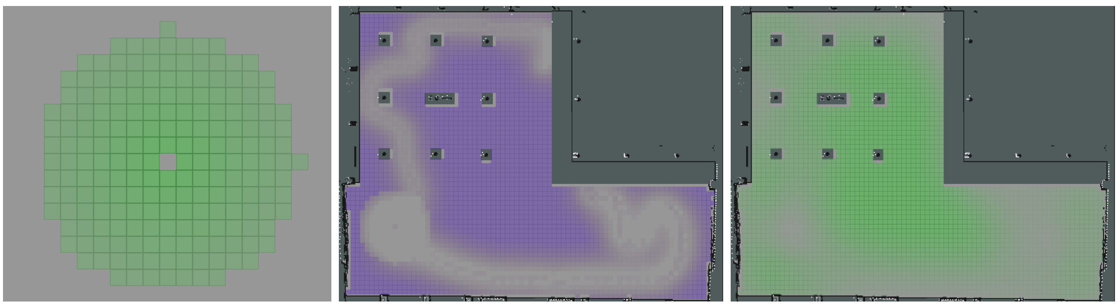


Figure 5.2: **Search Social Reward Sources.** From left to right: a) Individual search social reward source proportional to the probability of detecting the object in its center area. b) Visualization of the object location probability in an intermediate timestamp of a two-agent collaborative search. c) Additive visualization of the rewards generated by all the search reward sources in the previous timestamp.

6. User Studies & Experimentation

We designed two experiments to validate both the knowledge representation model and the goodness of the collaborative search task design.

6.1 SRS User Study

Some SRS model virtues are sustained over the assumption humans can easily comprehend and use SRSs. In open worlds, humans may have little to no time to train in the usage of robotic systems. We did a user experiment to test SRSs usability to build shared knowledge and diminish the adaptation process time.

A total of 20 non-trained volunteers participated in the experiment, with ages between 17 and 70 (mean: 30.65 std: 16.43). The educational level of the participants, as well as their fields of study, were diverse. On a scale of 1 (None) to 7 (Expert) their average self-evaluated knowledge in robotics was 1.7 (std: 1.26). No one had any experience using the framework, neither were they given the chance to practice.

Given the environments in Fig. 6.1, participants were told that the robot had either lost sensory capabilities or never had them and asked to indicate to it how to properly reach the goal while avoiding hazards. Possible interactions given to them were to tell the robot to avoid a zone, pass through a specific place, or go to some point. These instructions were modelled as cylindrical SRSs with boundary referenced Gaussian decay, being respectively a $constant$ repulsor and two kinds of $consumable$ attractors, one collectable over the whole robot path and another only meaningful when matched with the robot final goal. The three sources where correspondingly described as the instructions: "avoid this zone", "pass through this point" and "go to this place".

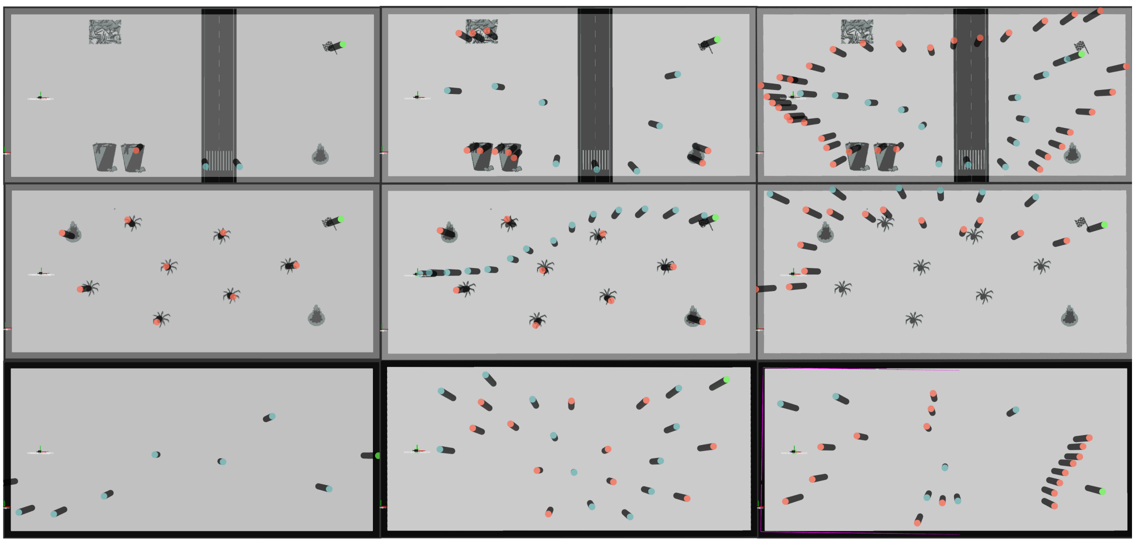


Figure 6.1: **User Study Environments.** Designs build by the participants in the user study. They were entitled to use three instructions: avoid this place (red cylinders), pass through this place (blue cylinders) and go to this place (green cylinders). Scenarios and corresponding indications given to the participants were, following rows from top to bottom. a) Crossroad: Guide the robot to reach the objective (flag) avoiding the objects and crossing the road through the crossroad. b) Spiders: Guide the robot to reach the objective however you feel fitter. c) Free Space: Imagine a trajectory and try to make the robot reproduce it.

The SRSs space was explored through the sampling-based approach explained in section 4. The simulation is run over gazebo and all navigation systems are integrated with the Robot Operating System (ROS). In all tests, attractive sources are erased when reached by the robot. Each of the participants was only permitted to solve each of the different environments once. After each test, participants were surveyed about whether the robot had planned what they expected and whether they found the planned path to be reasonable (sound according to the instructions given). Both questions were answered on a linear scale from 1 (not at all) to 7 (completely), results can be seen in Table 6.1.

Table 6.1: User Study Survey

	Crosswalk		**Spiders**		**Free**	
Environments	mean	std	mean	std	mean	std
Expected Path	5.75	1.68	6.6	0.68	5.55	1.53
Reasonable Path	6.2	1.39	6.75	0.44	5.85	1.66

Implementation designs were highly variable, yet all reached the given goals. Though designs in Fig. 6.1 show complete descriptions of the task, some participants decided to dynamically direct the robot through partial subgoals. Only three participants failed to completely avoid the obstacles in the first test, while all did it in the second. In most cases, differences between the first and the second questions on each episode were attributed by the participants to a personal error. Some participants grading lower the third test complained about the restrictiveness of their interactions, in particular not including larger repulsors and being unable to assign order relations in the attractors.

We conclude humans and robots achieved a shared task representation either on the first or the second test. Taking into account the test results, intuitiveness of the model is also assumed given the participants' background diversity, their previous knowledge and the absence of training.

6.2 Collaborative Search

We validate our model using the BRL map from the *Barcelona Robot Lab Dataset*[1], where three different locations are chosen as the search team origin (Fig. 6.2.a). The explorable area is discretised and all obstacles in the scene are assumed to block both the view of the robot and the human. We tested the model in a two-agent simulated environment, where human participants

[1] http://www.iri.upc.edu/research/webprojects/pau/datasets/BRL/

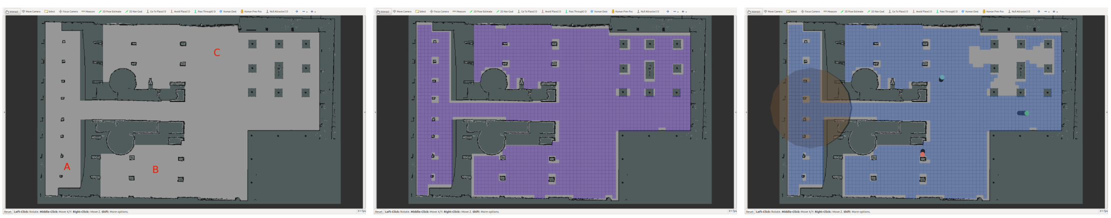


Figure 6.2: **Collaborative Search Experiments.** From left to right: a) BRL map and search team origins. b) Robot perceived exploration progress and visual feedback of the communication instructions given to the human. Three general instructions, “go to this place” (green cylinder), “pass through this place” (blue cylinder) and “avoid this place” (red cylinder); and 2 task-related informative messages, “I'm going to this place” (brown area) and “I've already been here” (perceived explored area at the top right zone of the map).

teleoperate an avatar that collaborates with the robot in the search process. In this first approach, we assumed an absolute detection capability until some distance threshold for both agents. This serves the purpose of simplifying the environment complexity and ease human comprehension of the task under simulation.

First, to establish a baseline, we studied human and robot individual search performance. After that, we tested three different communication levels to evaluate the model. In the first one, the human was shown the exploration progress and the robot location. In the second, the perceived exploration progress and his current planned path were added and, during the third experiment, the human could communicate with the robot through 5 instructions (Fig. 6.2.c).

A total of 12 volunteers participated in the experiment, with ages between 15 and 34 (mean: 26.4 std: 5.2). On a scale of 1 (None) to 7 (Expert) their average self-evaluated knowledge in robotics was 4.8 (std: 1.6). No one could practise using the framework, neither had any previous experience using it. Each of them participated in three of the different experimental setups involving humans, doing 3 or 6 episodes on each one equally distributed among the different origins. Additionally, participants were surveyed after each communication level setup whether they perceived robot plan as efficient and how much did they change their plans due to the robot actions. Both questions were answered on a linear scale from 1 (not at all) to 7 (completely).

Both the speed of the robot and the human were limited during all the experiments. Participants controlled the simulation through a *PlayStation 3 Dualshock 3 Wireless Controller* and could move at a maximum velocity of 1 m/s. The robot, the virtual model of a luggage transporter mounted on a *Pioneer P3-DX* base, had a maximum linear speed of 0.7 m/s, being it the nominal maximum velocity of the real robot. Human and robot mean speed along all the simulations were 0.83 m/s and 0.53 m/s, respectively.

6.2.1 Results

A summary plot of the collaborative search experiments is shown in Figure 6.3. As we can observe, origin selection proves to have a strong influence in the search progress dynamics. Additionally, we can perceive correlations between the human and the robot search progress shape, suggesting their search policies are alike.

Episodes beginning in B have the biggest robot contribution and robot behaviour shows greater variability when beginning in origin A, presumably due to the presence of two major bifurcations. Consistency in the collaborative search with communication dataset suggests that human users either instructed the robot where to go or implicitly conditioned its choice by providing it with information. As a matter of fact, all the participants preferred the robot to take the hallway while they explored the remaining area at their side. Moreover, most of them enforced this behaviour through direct orders, while the usage of the task-related informative messages was relegated only to the right part of the map.

Except for the late-stage search progress when beginning in origin B, all three collaborative models surpassed both the individual human and robot baselines. However, in terms of search progress, neither of the three is proven to be significantly better than the others. Possibly, the information given to the human on the first collaborative setup might be too extense. These results encourage further experiments conveying even less information to the human. Besides, we judge that the adaptation capabilities of the human, as well as their superior movement capabilities, made up for the lack of communication.

On the other hand, including human to robot communication greatly decreases the number of situations where the human is forced to adapt to the robot. Moreover, it seems to improve the

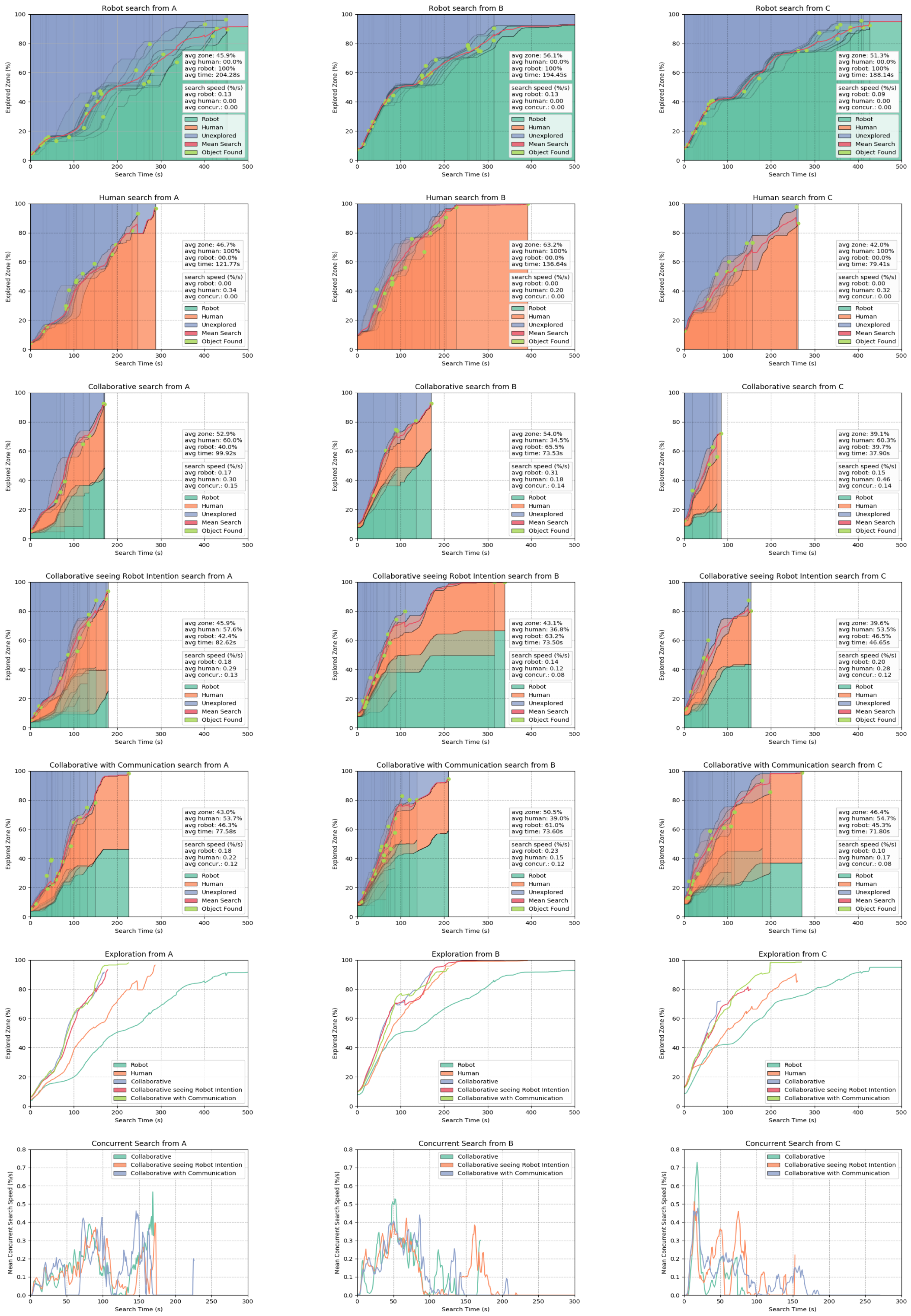


Figure 6.3: **Human-Robot Collaborative Search Experiments.** From left to right: episodes beginning at origins A, B and C. From top to bottom: Mean exploration in the 5 setups (robot individual search, human individual search, collaborative search, collaborative search seeing robot intention, collaborative search including human to robot communication) summarized comparison and concurrent activity.

human perception of the robot efficiency (Fig. 6.4). Even though in the second communication level the human had a broader perception of the robot intention, this might have enhanced conflict situations between the human-perceived robot plan and their own.

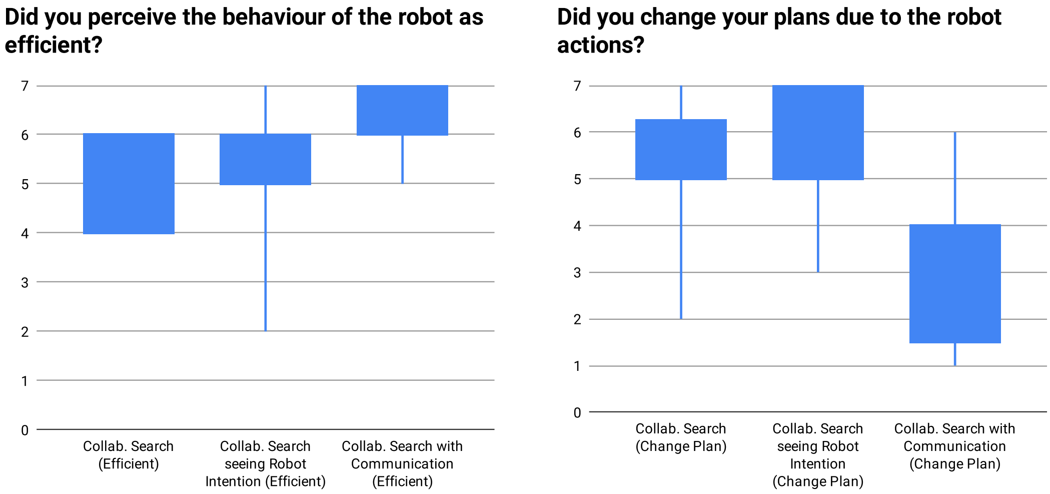


Figure 6.4: **Collaborative Search Survey.** Participants where asked whether they percieved the robot behaviour task-efficient and whether they had to change plans due to the robot actions. They answered both questions on a scale from 1 (Not at all) to 7 (Completelly).

7. Derived Projects

The previous first approach provides completeness (the robot will eventually find the object even in the absence of human contribution), performs a successful goal selection and achieves a task time reduction through human-robot concurrent activity. However, all said and done, it can be vastly improved. The performance benefits of collaboration are built upon the human adaptive capability and, at the end of the day, the human-robot pairs are interacting in a simulated environment.

7.1 Human Search Dataset

In the presented experiments, human and robot detection models have been simplified to distance thresholds. It is important to note that, in these, participants' knowledge about the task progression is limited to the interface's visual representation. A stochastic model, such as the one shown in Fig. 5.2, may be more realistic and informative for the robot, but humans have a strongly biased perception over probabilities. On one hand, most collaborative interactions still take place with a simple detection model, while on the other, episode length and human biases could generate frustration and a feeling of unfairness.

Nevertheless, though previous premises hold for a simulated environment, they are not applicable in the real world. If aiming to achieve meaningful collaboration in the physical world, a precise detection model for humans is needed. Moreover, in forthcoming approaches, the robot is expected to generate plans that take into consideration the predicted behaviour of the human (the presented model only contemplates the current agent contribution and explicit communication). Unfortunately, we lack knowledge over human strategic approach of this task and their detection capabilities. Besides, the latter is strongly variable, being affected by the searched object shape, colour and texture, the environment colours and light and the fatigue, between others.

To sum up, before attempting real-life human-robot collaborative search, we should collect information about single human and human-human collaborative search. In this section, we detail the characteristics of our new Human Search Dataset.

7.1.1 Data Acquisition

All the data have been collected in the Barcelona Robot Lab, whose map was used in the previous experiments. There, voluntary participants were asked to search for an object either alone or with a companion. In the latter cases, they were not allowed to convey information to one another, though they could observe the other's actions. Each episode continued until the discovery of the object or the participant's capitulation. For consistency and repeatability, all participants were asked to search for the same target, a green Parcheesi piece. We recorded all their movements through two 2D lasers distributed in the search space.

The final dataset is constituted by 52 trajectories, 18 single-person and 17 two-person episodes (Fig. 7.1). The data was recorded from 25 participants and collected in the spawn of 3 days. The object was found in 28 out of 35 experiments and, in total, all the episodes sum up to 2,5 hours of search activity.

7.1.2 Object Detection

To build a human detection model, we evaluate the trajectories in samples of $\Delta t = 0.1s$. A visualization of the different episodes' object location in the human reference may be observed in Fig. 7.2, as well as the data distribution over radial distance. Most detections occur on the participants' front, all corresponding to direct frontal approaches to the object. Only one instance happens to be located behind the person. It should be noted that, as they were looking for a small object, participants were most of the time looking downwards and constantly moving their head and visual focus, hence detections at the participant sides or back are possible. Actually, the higher density of object detections in the person's front may be more related to the person movement than visual its field of view (observe that most detected object trajectories in the human reference come straight from the front).

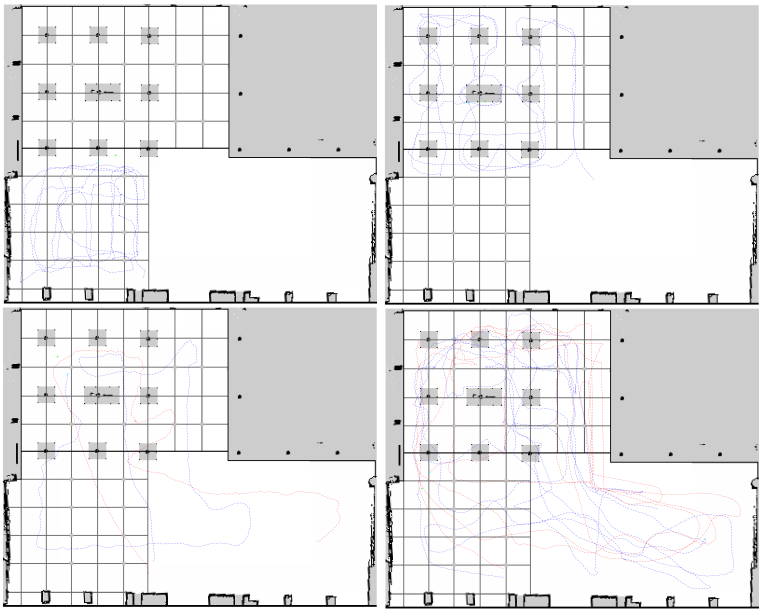


Figure 7.1: **Trajectory examples.** Trajectories corresponding to two single-person searches, one in a free zone and another in one with trees, and two two-person searches.

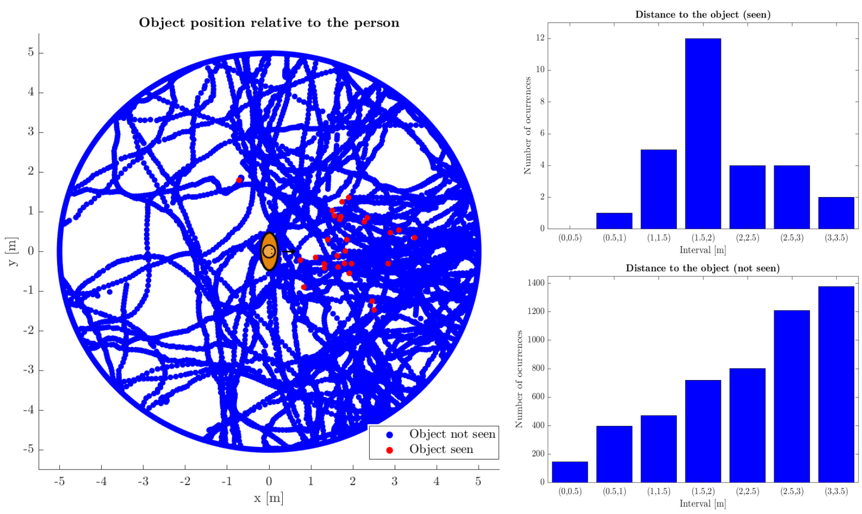


Figure 7.2: **Detection Data.** The left image displays all undetected (blue) and detected (red) object positions in the human reference (person front to the right). Histograms at the right show sample density over distance of detected (up) and undetected (down) instances.

7.2 Multi-agent MCTS

Good collaborative planning should integrate the teammates' contribution to the task. The previous first approach relies strongly on human adaptability to achieve fruitful cooperation. While the plan is constructed over the current task progress knowledge, the companions' position and potential future contribution are not taken into account.

We propose the usage of a multi-agent planner in human-robot collaborative settings. Building a shared plan for all the team members, even if not explicitly communicated, provides knowledge on each agent probable contribution. Using the insight from multi-robot and multi-agent literature, we aim to obtain both a team-aware motion plan for the robot and a task-biased movement prediction of the other agents.

What's more, a multi-agent planner can make full use of Social Reward Sources model. Exploiting sources' target sets, one may easily define shared goals and build agents' mental models. All the more so, in this setting team members themselves may be defined as sources of reward to model inter-agent dependencies.

Inspired by some recent multi-agent works [14, 10], we have decided to use a Monte Carlo Tree Search approach for human-robot settings. To deal with the exponential increase of the search space, we build the collaborative plan in two phases: the generation of a restricted action set and the collaborative plan expansion over it.

7.2.1 Glossary

$a \in \{1,2,...,A\}$	Agent a
$s \in S$	MCTS state s
$\mathcal{X}^a$	Feasible action set of agent a
$\hat{\mathcal{X}}^a \subset \mathcal{X}^a$	Restricted set of $\mathcal{X}^a$
$\boldsymbol{x} := \{\boldsymbol{x}^1,...,\boldsymbol{x}^A\}, \boldsymbol{x} \in \mathcal{X}$	Team action sequence
$\boldsymbol{x}^a = (x^a_{n_1}, x^a_{n_2}..), \boldsymbol{x}^a \in \mathcal{X}^a$	Agent a action sequence
$(t^a_{n_1}, t^a_{n_2}...)$	Agent a action times
$(c^a_{n_1}, c^a_{n_2}...)$	Agent a action costs
$q(\boldsymbol{x})$	Probability of $\boldsymbol{x}$ over $\hat{\mathcal{X}}$ (restricted team action set)
$q^a(\boldsymbol{x}^a)$	Probability of $\boldsymbol{x}^a$ over $\mathcal{X}^a$
$g(\boldsymbol{x})$	Global objective function
$\psi^\tau \in \Psi$	Social reward source with targets τ
R^a_A	Actions' reward
R^a_E	Environment reward
R_T	Collaborative task reward
R_I	Team interaction reward

7.2.2 Search Space

In this MCTS implementation, we use the previously presented RRT to generate a feasible restricted set of each agent's possible paths $\hat{\mathcal{X}}^a$. These paths can be seen as heterogeneous action sequences, each one leading to a different goal.

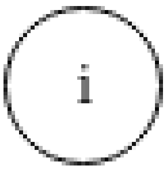
In Best et al. [10], they build a multi-agent MCTS planner over a common PRM. Their agents, however, are assumed to have the same mobile and planning capabilities, a premise that does not hold in human-robot settings. Using individual RRTs, each agent tree may be expanded with different restrictions and precision. For example, one agent may need to avoid a specific spot (e.g.

hazards) while others can move through safely. Likewise, a distant agent tree may be generated with large node distances, obtaining information about its contribution but not about its movement dynamics, to balance the search space dimensionality.

7.2.2.1 Agent Action Set

Each agent action is assumed to be a movement action and represented by an RRT node i. Movement actions are defined by their origin, the node's parent location, their goal, their own location, and the completion time. Additionally, each action node can store a distribution probability over its children electability on a satisfactory shared plan.

Every RRT node i can have an unbounded number of children $ch(i)$, but the number of actions eligible after each agent action is bounded to N_a (this is further developed in section 7.2.3.1)

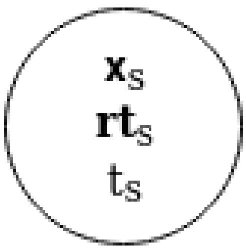


$\vec{p_i}^a$: target search space location of x_i^a
t_i^a: action x_i^a time
c_i^a: cost of action x_i^a
q_i^a: probability distribution over node i's children

7.2.2.2 MCTS State

MCTS states are defined over the agents' action RRT trees. Each state s is formed by a list of ongoing agent actions $\boldsymbol{x}_s$ and each one's remaining time to finish $\boldsymbol{rt}_s$, as well as the current collaborative plan time of the state.

Each MCTS state s_k can have a determinate number of children states $ch(s_k)$. The number of children states is bounded by $|ch(s_k)| \leqslant \prod_a^{(a_i|rt_{s_k}^{a_i}=0)} ch(x_{s_k}^a)$. In other words, each MCTS state can only have as many successors as the existent possible combinations of finished tasks' eligible children. Moreover, a new state should be generated after each agent task is finished. More insight about this restrictions can be found in section 7.2.3.1.



$\boldsymbol{x}_s = \{x_s^0, x_s^1, \ldots, x_s^A\}$: ongoing actions on time t_s
$\boldsymbol{rt}_s = \{rt_s^1, rt_s^2, \ldots, rt_s^A\}$: remaining time to finish $\boldsymbol{x}_s$
t_s: inferred time for state s

7.2.3 Collaborative Planning

We aim to build a collaborative navigation plan to tackle shared tasks over heterogeneous action sets with variable time horizons. To do so, the MCTS planner should ensure temporal coherence in the tree expansion, deal with coexistence of agent action sequences with different temporal length and provide a feasible reward propagation mechanism to deal with a dynamic environment.

7.2.3.1 MCTS expansion

First, an individual agent reward upper bound is calculated over each agent RRT. Rewards for each agent a action sequences $\boldsymbol{x}^a \in \mathcal{X}^a$ are updated using the method explained in Section 4.2. Then, all these rewards are back-propagated and each tree node i stores the maximum attainable individual reward from it. For each, only the N_a children nodes with the highest upper bound are considered, further pruning the search tree, and a selection probability q_i^a weighted by those bounds is assigned to them.

The multi-agent plan is expanded from the root node s_0, which is constituted by the root actions of each agent (0-time actions). Every iteration m, one of the expandable states s_m in the collaborative plan tree is randomly selected. From it, a chain of future states is continuously simulated until a final state is reached. Each new simulated state is sampled from a probability distribution q defined by each individual agent action node probability distribution q^a.

An example of an ongoing collaborative plan expansion may be found in Fig. 7.3.

Additionally, at a latter stage, we added some predefined preliminary expansion to the MCTS tree. The previous model eventually finds good plans, but the exponential nature of the search space may prove problematic in time constrained applications. To ensure the viability of early plans we define a set of end state candidates for the tree. As a first step, we collect potentially rewarding

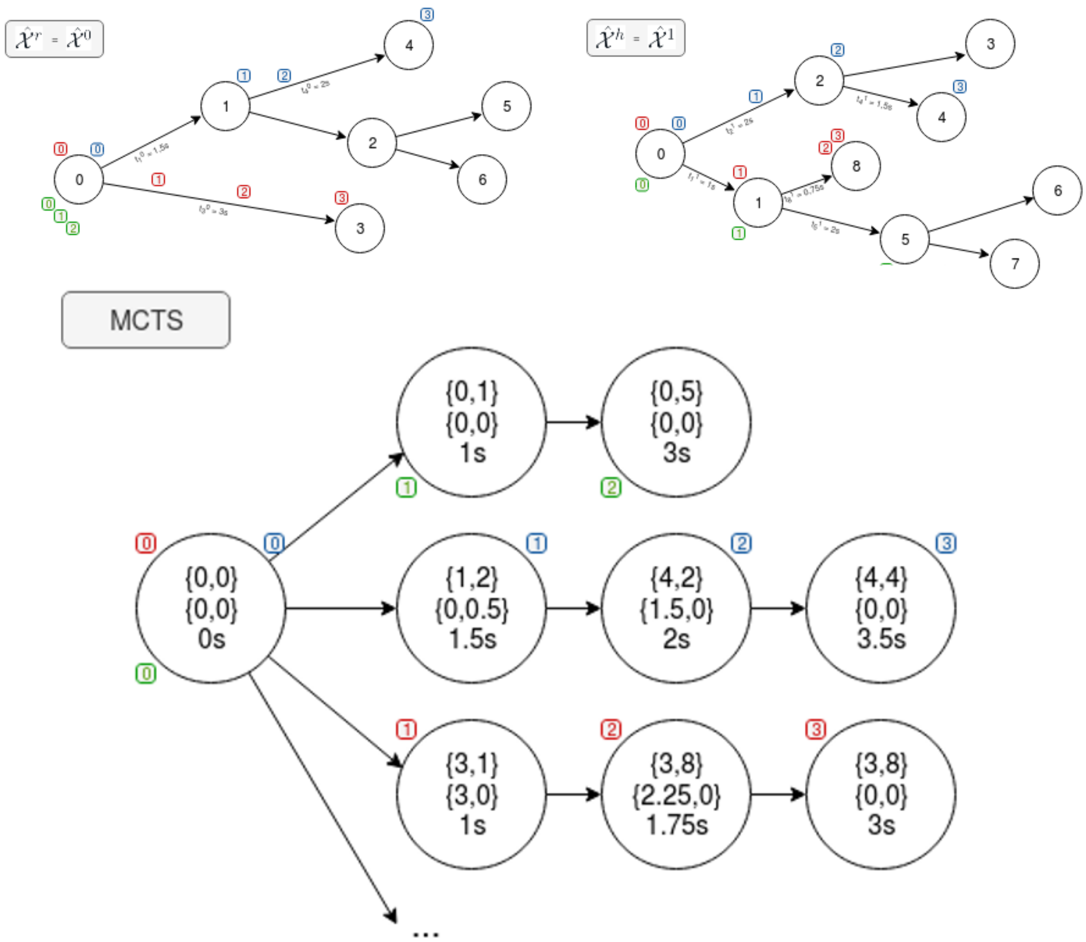


Figure 7.3: **MCTS Tree expansion.** Example of a collaborative plan expansion for a two-member team. The green, blue and red markers illustrate three possible plans and each agent RRT mapping of their MCTS nodes.

goals from the individual agents' action sequences and combine them to form goal states. After that, the global team plans leading to these goal states are generated. Note that such connection is deterministic if agents are assumed to continually take consequent actions Temporal variations of these paths can be generated *a posteriori* during the standard expansion algorithm.

7.2.3.2 Reward Propagation

We may define the collaborative plan objective function $g(\boldsymbol{x})$ as the additive combination of all the rewards influencing the team. This includes the rewards related to each agent actions cost and perceived influence of the environment, the shared task and team interaction.

$$g(\boldsymbol{x}) = \sum_{a}^{\{1,..,A\}} (R_A^a + R_E^a) + R_T + R_I$$

being

$$R_A^a(\boldsymbol{x}^a) = -\sum_{x_j^a \in \boldsymbol{x}^a} c_j^a \qquad R_T(\boldsymbol{x}) = \sum_{s_k}^{S(\boldsymbol{x})} r_t(\Psi_t^\tau, \boldsymbol{x}_{s_{0\to k}}, t_{s_k}, \Delta t_{s_k})$$

$$R_E^a(\boldsymbol{x}^a) = \sum_{s_k}^{S(\boldsymbol{x}^a)} r_e(\Psi_e^a, x_{s_k}^a, t_{s_k}, \Delta t_{s_k}) \qquad R_I(\boldsymbol{x}) = \sum_{s_k}^{S(\boldsymbol{x})} \sum_{a}^{\{1,..,A\}} r_i(\Psi_{i_a}^\tau, \boldsymbol{x}_{s_{0\to k}}^{\tau\cup a}, t_{s_k}, \Delta t_{s_k})$$

where $S(\boldsymbol{x})$ is the set of MCTS states defined by the action sequences $\boldsymbol{x}$. $r_e(\Psi_e, x^a, t, \Delta t)$ is the reward generated by the environmental sources set Ψ_e to target agent a while performing action x^a during a period of Δt initiated at time t. Likewise, r_t and r_i are the rewards generated by sources Ψ_t^τ and $\Psi_{i_a}^\tau$, given the action sequences in $\boldsymbol{x}_{s_{0\to k}}$ (action sequences that generate the tree branch connecting the initial state s_0 and the state s_k). Finally,

$$rt_s^a = \begin{cases} rt_s^a - \Delta t_s & \text{if } x_s^a = x_{p(s)}^a \\ t_s^a - \Delta t_s & \text{otherwise} \end{cases} \qquad \Delta t_s = t_s - t_{p(s)}$$

where $p(s)$ is the parent state of s.

7.2.4 Evaluation

The images in Fig. 7.4 are presented to provide some qualitative evaluation of the model. In them, a team of three agents distributed across the scene is given three shared goals to fulfill, if possible. We are shown the plan built by one of the agents using the multi-agent MCTS model over the SRS representation of four different environments.

Unfortunately, we were unable to provide any experimental evaluation of the model due to the present quarantine imposed by the COVID-19 world crisis.

Figure 7.4: **Multi-agent MCTS.** Three-agent examples of the presented MCTS method. From left to right, up to bottom: a) Global plan fulfilling three shared goals. b) Global plan maximizing global rewards, even at the expense of increasing individual agent's effort. c) One of the shared goals is inaccessible due to some hazard. In consequence, only two agents are expected to move towards the remaining goals. d) Now, there are two hazards in the scene, but they only affect one of the agents (green). The planner adapts to keep fulfilling the three shared goals.

8. Impact & Sustainability

This thesis has been developed within the mobile robotics laboratory at the *Institut de Robòtica i Informàtica Industrial (IRI CSIC-UPC)*[1]. The work has been supported under projects ColRobTransp (DPI2016-78957-RAEI/FEDER EU), TERRINet (H2020-INFRAIA-2017-1-two-stage-730994) and by the Spanish State Research Agency through the Maria de Maeztu Seal of Excellence to IRI (MDM-2016-0656). The Social Reward Sources model and a first collaborative search approach have been published at the *Fourth Iberian Robotics Conference* (ROBOT2019)[2] [24] and, hopefully, the multiple agent SRS representation and MCTS extension is going to be submitted as a Journal paper in the following months.

Due to this thesis focus in basic research, it has produced no major environmental or economical impact. Nevertheless, it is worth pointing out its effects and discussing the social impact of the global academic effort to which this work tries to contribute.

Environmentally speaking, the major hazard of this project is the related energy consumption. Only computer usage during this period has been computed to be equivalent to the emission of 389 Kg of CO_2. Other sources of pollution may include heating, light and water, all consumed in shared spaces. To assume an upper bound, this project estimated environmental impact might be equivalent to the emission of 550 Kg CO_2 into the atmosphere.

The social impact of this work is limited to its contribution to the academic research effort to push forward human knowledge and technology. As a whole, however, this process creates immense changes in society. Being more specific, the field of collaborative robotics is pushing forward the robots' capability to deal with open-world scenarios and exploring the potential of human-robot collaboration. In a context where new world challenges arise, robotics may be the answer to a lot of our problems. Robotics research and development is bringing major changes to the world as we know it, either from the social, economic or environmental perspective. May them be good or bad, only time will tell.

[1] www.iri.upc.edu

[2] https://web.fe.up.pt/ robot2019/

8.1 Budget

The estimated cost of this project sums up to 22.427,10 €. This includes 12 months of the full-time salary of a graduate (37,5 h/week), an average of 0,5 hours a week of a lab. technician and the supervisor dedication. Equally, it adds an estimation of the amortization and energetic cost of the project development. Details are depicted in Table 8.1.

Table 8.1: Budget

Concept	**Cost**			
Personnel	**Dedication [h]**	**Cost [€/h]**		**Total [€]**
Graduate Researcher	2025	10,20		20.655,00
Lab. Technician	27	14,80		399,60
PhD Supervisor	50	18,50		925,00
Equipment	**Price [€]**	**Amortization [years]**	**Usage Time [years]**	**Total [€]**
Computer	1400	5	1	280,00
Sensors	500	5	0,055	5,50
Others	**Usage [h]**	**Cost [€/h]**		**Total [€]**
Energy Consumption	2025	0,08		16,00
Total				**22.427,10 €**

Conclusions

In this thesis, we present a model for unified knowledge representation for human-robot collaboration, develop a motion planning approach that plans over this representation and validate the whole model implementing the human-robot collaborative search task.

The knowledge representation model, namely the Social Reward Sources model, is presented as a useful framework to design human-robot collaboration. The user study proves that humans can relate to this model and even design successful robot behaviours on the first try.

The motion planning approach is thoroughly validated in the simulated collaborative search testbed. In this setting, the presented task model outperforms the individual search baseline. Moreover, human to robot communication is proven to have a major impact in human perception of human-robot collaborative tasks, while performance might not be significantly influenced due to the human adaptation capabilities.

That being said, major weaknesses of the model, such as ongoing barriers to real-world implementation and strong dependence on human adaption capabilities, are pointed out. Two spin-off projects aiming to tackle this issue are presented, the compilation of a human-human collaborative search dataset and the development of a multi-agent motion planning approach to further improve the robot awareness of the collaborative task. Unfortunately, the experimental campaign aimed to validate this last approach was cancelled due to the current health crisis.

In conclusion, this thesis presents and validates a first approach to a state of the art problem. Also, it identifies current major problems in the followed approach and proposes a second theoretical model that engages them.

In future work, the models will be tested in a real-world environment. Additionally, the multi-agent approach can be further improved adopting variations present in the literature, such as node selection policies. We also envision the usage of the SRS model to express human preferences.

Acknowledgment

I have to start thanking my beloved family, their example, effort and support made possible that I'm here writing this thesis, after so many years focused on study. I hope to return, at least, a little portion of your dedication.

I'm also inevitably indebted to my fourth floor friends. Their influence in my life since the start of my academic journey have enriched both my knowledge and my life. I hope the mutual support of the past years proves to everlasting in our lives. Equally, my dearest neighbour, thanks. Your continuous support and positivity, specially during this quarantine, pushed me to reach this milestone. Finally, a special thanks to all the others who take part in the truly important pillars of my life. Albert, Ivan, Joan Ramon, members of Envers, beloved Senet friends and many others, it is the happiness you created what served as both a motor and a pillow when engaging this kind of challenges and projects.

Finally, I want to express my gratitue to all the people who invested time in the development of this work. Anañis Garrell, Pablo Jiménez and Alberto Sanfeliu, thanks for your dedication sharing my path, directing my efforts and checking my work. A special thank you to Albert Bhagwan, for his contribution in the dataset collection and processing, and Fernando Herrero, Sergi Hernández and Alejandro Lopez, for their patient and continuous support as laboratory technicians. A final mention to all my IRI colleagues contributing to generate a welcoming environment, my grateful recognition.

Bibliography

[1] Admoni, H., Scassellati, B.: Social eye gaze in human-robot interaction: a review. Journal of Human-Robot Interaction 6(1), 25–63 (2017) 6

[2] Albers, S., Kursawe, K., Schuierer, S.: Exploring unknown environments with obstacles. Algorithmica 32(1), 123–143 (2002) 10

[3] Aronson, R.M., Santini, T., Kübler, T.C., Kasneci, E., Srinivasa, S., Admoni, H.: Eye-hand behavior in human-robot shared manipulation. In: Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction. pp. 4–13. ACM (2018) 6

[4] Atmaca, S., Sebanz, N., Knoblich, G.: The joint flanker effect: sharing tasks with real and imagined co-actors. Experimental brain research 211(3-4), 371–385 (2011) 4

[5] Atmaca, S., Sebanz, N., Prinz, W., Knoblich, G.: Action co-representation: the joint snarc effect. Social neuroscience 3(3-4), 410–420 (2008) 4

[6] Bai, X., Yan, W., Ge, S.S., Cao, M.: An integrated multi-population genetic algorithm for multi-vehicle task assignment in a drift field. Information Sciences 453, 227–238 (2018) 10

[7] Balakirsky, S., Schlenoff, C., Rama Fiorini, S., Redfield, S., Barreto, M., Nakawala, H., Carbonera, J.L., Soldatova, L., Bermejo-Alonso, J., Maikore, F., et al.: Towards a robot task ontology standard. In: ASME 2017 12th International Manufacturing Science and Engineering Conference collocated with the JSME/ASME 2017 6th International Conference on Materials and Processing. American Society of Mechanical Engineers Digital Collection (2017) 5

[8] Bera, A., Randhavane, T., Prinja, R., Kapsaskis, K., Wang, A., Gray, K., Manocha, D.: The emotionally intelligent robot: Improving social navigation in crowded environments. arXiv preprint arXiv:1903.03217 (2019) 8

[9] Bera, A., Randhavane, T., Prinja, R., Manocha, D.: Sociosense: Robot navigation amongst pedestrians with social and psychological constraints. In: 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 7018–7025. IEEE (2017) 8

[10] Best, G., Cliff, O.M., Patten, T., Mettu, R.R., Fitch, R.: Dec-mcts: Decentralized planning for multi-robot active perception. The International Journal of Robotics Research 38(2-3), 316–337 (2019) 10, 31

[11] Bevilacqua, P., Frego, M., Bertolazzi, E., Fontanelli, D., Palopoli, L., Biral, F.: Path planning maximising human comfort for assistive robots. In: 2016 IEEE Conference on Control Applications (CCA). pp. 1421–1427. IEEE (2016) 8

[12] Bisk, Y., Yuret, D., Marcu, D.: Natural language communication with robots. In: Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. pp. 751–761 (2016) 6

[13] Bratman, M.E.: Shared cooperative activity. The philosophical review 101(2), 327–341 (1992) 4

[14] Browne, C.B., Powley, E., Whitehouse, D., Lucas, S.M., Cowling, P.I., Rohlfshagen, P., Tavener, S., Perez, D., Samothrakis, S., Colton, S.: A survey of monte carlo tree search methods. IEEE Transactions on Computational Intelligence and AI in games 4(1), 1–43 (2012) 10, 31

[15] Carlson, T., Demiris, Y.: Collaborative control for a robotic wheelchair: evaluation of performance, attention, and workload. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 42(3), 876–888 (2012) 8

[16] Castiello, U.: Understanding other people's actions: intention and attention. Journal of Experimental Psychology: Human Perception and Performance 29(2), 416 (2003) 4

[17] Chakraborti, T., Kulkarni, A., Sreedharan, S., Smith, D.E., Kambhampati, S.: Explicability? legibility? predictability? transparency? privacy? security? the emerging landscape of interpretable agent behavior. In: Proceedings of the International Conference on Automated Planning and Scheduling. vol. 29, pp. 86–96 (2019) 9

[18] Chakraborti, T., Sreedharan, S., Grover, S., Kambhampati, S.: Plan explanations as model reconciliation–an empirical study. arXiv preprint arXiv:1802.01013 (2018) 7

[19] Chen, C., Liu, Y., Kreiss, S., Alahi, A.: Crowd-robot interaction: Crowd-aware robot navigation with attention-based deep reinforcement learning. arXiv preprint arXiv:1809.08835 (2018) 8

[20] Chiang, H.T.L., Faust, A., Fiser, M., Francis, A.: Learning navigation behaviors end-to-end with autorl. IEEE Robotics and Automation Letters 4(2), 2007–2014 (2019) 8

[21] Clark, H.H., Schreuder, R., Buttrick, S.: Common ground at the understanding of demonstrative reference. Journal of verbal learning and verbal behavior 22(2), 245–258 (1983) 4

[22] Clodic, A., Pacherie, E., Alami, R., Chatila, R.: Key elements for human-robot joint action. In: Sociality and Normativity for Robots, pp. 159–177. Springer (2017) 6

[23] Costa, M.: Interpersonal distances in group walking. Journal of Nonverbal Behavior 34(1), 15–26 (2010) 5

[24] Dalmasso, M., Garrell, A., Jimenez, P., Sanfeliu, A.: Human-robot collaborative navigation search using social reward sources. arXiv preprint arXiv:1909.04768 (2019) 37

[25] Dautenhahn, K., Walters, M., Woods, S., Koay, K.L., Nehaniv, C.L., Sisbot, A., Alami, R., Siméon, T.: How may i serve you?: a robot companion approaching a seated person in a helping context. In: Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction. pp. 172–179. ACM (2006) 8

[26] Devin, S., Alami, R.: An implemented theory of mind to improve human-robot shared plans execution. In: 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI). pp. 319–326. IEEE (2016) 7

[27] Dolk, T., Hommel, B., Colzato, L.S., Schütz-Bosbach, S., Prinz, W., Liepelt, R.: How "social" is the social simon effect? Frontiers in Psychology 2, 84 (2011) 4

[28] Dolk, T., Hommel, B., Prinz, W., Liepelt, R.: The (not so) social simon effect: a referential coding account. Journal of Experimental Psychology: Human Perception and Performance 39(5), 1248 (2013) 4

[29] Dragan, A.D., Bauman, S., Forlizzi, J., Srinivasa, S.S.: Effects of robot motion on human-robot collaboration. In: Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction. pp. 51–58. ACM (2015) 8

[30] Elliot, A.J., Thrash, T.M.: Approach-avoidance motivation in personality: approach and avoidance temperaments and goals. Journal of personality and social psychology 82(5), 804 (2002) 3, 11

[31] Everett, M., Chen, Y.F., How, J.P.: Motion planning among dynamic, decision-making agents with deep reinforcement learning. In: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 3052–3059. IEEE (2018) 8

[32] Fadiga, L., Fogassi, L., Pavesi, G., Rizzolatti, G.: Motor facilitation during action observation: a magnetic stimulation study. Journal of neurophysiology 73(6), 2608–2611 (1995) 4

[33] Ferrer, G., Garrell, A., Sanfeliu, A.: Robot companion: A social-force based approach with human awareness-navigation in crowded environments. In: IEEE/RSJ International Conference on Intelligent Robots and Systems. pp. 1688–1694 (2013) 8

[34] Ferrer, G., Sanfeliu, A.: Bayesian human motion intentionality prediction in urban environments. Pattern Recognition Letters 44, 134–140 (2014) 5

[35] Ferrer, G., Zulueta, A.G., Cotarelo, F.H., Sanfeliu, A.: Robot social-aware navigation framework to accompany people walking side-by-side. Autonomous robots 41(4), 775–793 (2017) 8

[36] Fong, T., Thorpe, C., Baur, C.: Collaboration, dialogue, human-robot interaction. In: Robotics Research, pp. 255–266. Springer (2003) 8

[37] Freedy, A., DeVisser, E., Weltman, G., Coeyman, N.: Measurement of trust in human-robot collaboration. In: 2007 International Symposium on Collaborative Technologies and Systems. pp. 106–114. IEEE (2007) 9

[38] Gallese, V., Fadiga, L., Fogassi, L., Rizzolatti, G.: Action recognition in the premotor cortex. Brain 119(2), 593–609 (1996) 4

[39] Gallotti, M., Frith, C.D.: Social cognition in the we-mode. Trends in cognitive sciences 17(4), 160–165 (2013) 4

[40] Garrell, A., Garza-Elizondo, L., Villamizar, M., Herrero, F., Sanfeliu, A.: Aerial social force model: A new framework to accompany people using autonomous flying robots. In: 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2017, Vancouver, BC, Canada, September 24-28, 2017. pp. 7011–7017 (2017) 8

[41] Garrell, A., Sanfeliu, A.: Cooperative social robots to accompany groups of people. The International Journal of Robotics Research 31(13), 1675–1701 (2012) 8

[42] Goldhoorn, A., Garrell, A., Alquézar, R., Sanfeliu, A.: Searching and tracking people with cooperative mobile robots. Autonomous Robots 42(4), 739–759 (2018) 8

[43] González-Banos, H., Latombe, J.C.: Planning robot motions for range-image acquisition and automatic 3d model construction (1998) 10

[44] Greenwald, A.G.: A choice reaction time test of ideomotor theory. Journal of Experimental Psychology 86(1), 20 (1970) 3

[45] Haddad, S., Wu, M., Wei, H., Lam, S.K.: Situation-aware pedestrian trajectory prediction with spatio-temporal attention model. arXiv preprint arXiv:1902.05437 (2019) 5

[46] Hall, E.T., Birdwhistell, R.L., Bock, B., Bohannan, P., Diebold Jr, A.R., Durbin, M., Edmonson, M.S., Fischer, J., Hymes, D., Kimball, S.T., et al.: Proxemics [and comments and replies]. Current anthropology 9(2/3), 83–108 (1968) 4

[47] Hall, E.T.: The hidden dimension, vol. 609. Garden City, NY: Doubleday (1966) 4, 11

[48] Helbing, D., Farkas, I.J., Molnar, P., Vicsek, T.: Simulation of pedestrian crowds in normal and evacuation situations. Pedestrian and evacuation dynamics 21(2), 21–58 (2002) 5

[49] Helbing, D., Molnar, P.: Social force model for pedestrian dynamics. Physical review E 51(5), 4282 (1995) 3, 5, 11

[50] Hoffman, G.: Evaluating fluency in human-robot collaboration. IEEE Transactions on Human-Machine Systems (2019) 9

[51] Hoffman, G., Breazeal, C.: Collaboration in human-robot teams. In: AIAA 1st Intelligent Systems Technical Conference. p. 6434 (2004) 6

[52] Hommel, B., Colzato, L.S., Van Den Wildenberg, W.P.: How social are task representations? Psychological Science 20(7), 794–798 (2009) 4

[53] Hommel, B., Müsseler, J., Aschersleben, G., Prinz, W.: The theory of event coding (tec): A framework for perception and action planning. Behavioral and brain sciences 24(5), 849–878 (2001) 3, 11

[54] Hong, A., Igharoro, O., Liu, Y., Niroui, F., Nejat, G., Benhabib, B.: Investigating human-robot teams for learning-based semi-autonomous control in urban search and rescue environments. Journal of Intelligent & Robotic Systems 94(3-4), 669–686 (2019) 7

[55] Hu, J.S., Wang, J.J., Ho, D.M.: Design of sensing system and anticipative behavior for human following of mobile robots. IEEE Transactions on Industrial Electronics 61(4), 1916–1927 (2013) 8

[56] Huang, C.M., Andrist, S., Sauppé, A., Mutlu, B.: Using gaze patterns to predict task intent in collaboration. Frontiers in psychology 6, 1049 (2015) 4

[57] Jayawardena, C., Ardekani, I., et al.: A navigation model for side-by-side robotic wheelchairs for optimizing social comfort in crossing situations. Robotics and Autonomous Systems 100, 27–40 (2018) 8

[58] Johansson, A., Helbing, D., Shukla, P.K.: Specification of the social force pedestrian model by evolutionary adjustment to video tracking data. Advances in complex systems 10(supp02), 271–288 (2007) 5

[59] Johnson, M., Bradshaw, J.M., Feltovich, P.J., Jonker, C.M., Van Riemsdijk, M.B., Sierhuis, M.: Coactive design: Designing support for interdependence in joint activity. Journal of Human-Robot Interaction 3(1), 43–69 (2014) 7, 9, 12

[60] Jonker, C.M., Van Riemsdijk, B., Van De Kieft, I.C., Gini, M.: Towards measuring sharedness of team mental models by compositional means. In: AAAI Fall Symposium Series (2011) 7, 10

[61] Kanda, T., Glas, D.F., Shiomi, M., Hagita, N.: Abstracting people's trajectories for social robots to proactively approach customers. IEEE Transactions on Robotics 25(6), 1382–1396 (2009) 8

[62] Karamouzas, I., Skinner, B., Guy, S.J.: Universal power law governing pedestrian interactions. Physical review letters 113(23), 238701 (2014) 5

[63] Kasper, W.: Team monitoring and reporting for robot-assisted usar missions. In: International Symposium on Safety, Security, and Rescue Robotics (SSRR). pp. 246–251. IEEE (2016) 7

[64] Khambhaita, H., Alami, R.: Assessing the social criteria for human-robot collaborative navigation: A comparison of human-aware navigation planners. In: 2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN). pp. 1140–1145. IEEE (2017) 8

[65] Khambhaita, H., Alami, R.: Viewing robot navigation in human environment as a cooperative activity. arXiv preprint arXiv:1708.01267 (2017) 8

[66] Kitazawa, K., Fujiyama, T.: Pedestrian vision and collision avoidance behavior: Investigation of the information process space of pedestrians using an eye tracker. In: Pedestrian and evacuation dynamics 2008, pp. 95–108. Springer (2010) 5

[67] Kruijff, G.J.M., Janíček, M., Keshavdas, S., Larochelle, B., Zender, H., Smets, N.J., Mioch, T., Neerincx, M.A., Diggelen, J., Colas, F., et al.: Experience in system design for human-robot teaming in urban search and rescue. In: Field and Service Robotics. pp. 111–125. Springer (2014) 7

[68] Kruijff-Korbayová, I., Colas, F., Gianni, M., Pirri, F., de Greeff, J., Hindriks, K., Neerincx, M., Ögren, P., Svoboda, T., Worst, R.: Tradr project: Long-term human-robot teaming for robot assisted disaster response. KI-Künstliche Intelligenz 29(2), 193–201 (2015) 7

[69] Kruse, T., Kirsch, A., Khambhaita, H., Alami, R.: Evaluating directional cost models in navigation. In: Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction. pp. 350–357. ACM (2014) 8

[70] Lasota, P.A., Shah, J.A.: Analyzing the effects of human-aware motion planning on close-proximity human-robot collaboration. Human factors 57(1), 21–33 (2015) 8

[71] Li, M., Yang, W., Cai, Z., Yang, S., Wang, J.: Integrating decision sharing with prediction in decentralized planning for multi-agent coordination under uncertainty. In: Proceedings of the 28th International Joint Conference on Artificial Intelligence. pp. 450–456. AAAI Press (2019) 10

[72] Limerick, H., Coyle, D., Moore, J.W.: The experience of agency in human-computer interactions: a review. Frontiers in human neuroscience 8, 643 (2014) 4

[73] Lin, Z., Zhao, L., Yang, D., Qin, T., Liu, T.Y., Yang, G.: Distributional reward decomposition for reinforcement learning. In: Advances in Neural Information Processing Systems. pp. 6212–6221 (2019) 6, 11

[74] Liu, A.A., Su, Y.T., Nie, W.Z., Kankanhalli, M.: Hierarchical clustering multi-task learning for joint human action grouping and recognition. IEEE transactions on pattern analysis and machine intelligence 39(1), 102–114 (2016) 4

[75] Liu, H., Wang, L.: Gesture recognition for human-robot collaboration: A review. International Journal of Industrial Ergonomics 68, 355–367 (2018) 6

[76] Losey, D.P., McDonald, C.G., Battaglia, E., O'Malley, M.K.: A review of intent detection, arbitration, and communication aspects of shared control for physical human-robot interaction. Applied Mechanics Reviews 70(1), 010804 (2018) 6

[77] Luber, M., Spinello, L., Silva, J., Arras, K.O.: Socially-aware robot navigation: A learning approach. In: 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems. pp. 902–907. IEEE (2012) 5

[78] Marconi, L., Melchiorri, C., Beetz, M., Pangercic, D., Siegwart, R., Leutenegger, S., Carloni, R., Stramigioli, S., Bruyninckx, H., Doherty, P., et al.: The sherpa project: Smart collaboration between humans and ground-aerial robots for improving rescuing activities in alpine environments. In: International Symposium on Safety, Security, and Rescue Robotics (SSRR). pp. 1–4. IEEE (2012) 7

[79] Mavridis, N.: A review of verbal and non-verbal human-robot interactive communication. Robotics and Autonomous Systems 63, 22–35 (2015) 6

[80] Mead, R., Atrash, A., Matarić, M.J.: Automated proxemic feature extraction and behavior recognition: Applications in human-robot interaction. International Journal of Social Robotics 5(3), 367–378 (2013) 5

[81] Michael, J., Sebanz, N., Knoblich, G.: The sense of commitment: A minimal approach. Frontiers in Psychology 6, 1968 (2016) 4

[82] Mioch, T., Peeters, M.M., Neerincx, M.A.: Improving adaptive human-robot cooperation through work agreements. In: 27th International Symposium on Robot and Human Interactive Communication (RO-MAN). pp. 1105–1110. IEEE (2018) 7

[83] Moore, J.W.: What is the sense of agency and why does it matter? Frontiers in psychology 7, 1272 (2016) 4

[84] Morales, Y., Miyashita, T., Hagita, N.: Social robotic wheelchair centered on passenger and pedestrian comfort. Robotics and Autonomous Systems 87, 355–362 (2017) 8

[85] Moussaïd, M., Perozo, N., Garnier, S., Helbing, D., Theraulaz, G.: The walking behaviour of pedestrian social groups and its impact on crowd dynamics. PloS one 5(4), e10047 (2010) 5

[86] Neal, A., Ballard, T., Vancouver, J.B.: Dynamic self-regulation and multiple-goal pursuit. Annual Review of Organizational Psychology and Organizational Behavior 4, 401–423 (2017) 4, 11

[87] Nikolaidis, S., Nath, S., Procaccia, A.D., Srinivasa, S.: Game-theoretic modeling of human adaptation in human-robot collaboration. In: 2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI. pp. 323–331. IEEE (2017) 7

[88] Nunes, E., Manner, M., Mitiche, H., Gini, M.: A taxonomy for task allocation problems with temporal and ordering constraints. Robotics and Autonomous Systems 90, 55–70 (2017) 10

[89] Pacherie, E.: How does it feel to act together? Phenomenology and the cognitive sciences 13(1), 25–46 (2014) 4

[90] Pfeiffer, U., Schilbach, L., Timmermans, B., Jording, M., Bente, G., Vogeley, K.: Eyes on the mind: investigating the influence of gaze dynamics on the perception of others in real-time social interaction. Frontiers in psychology 3, 537 (2012) 4

[91] Rajpathak, D., Motta, E.: An ontological formalization of the planning task. In: International Conference on Formal Ontology in Information Systems (FOIS 2004). pp. 305–316 (2004) 5

[92] Ramírez, O.A.I., Khambhaita, H., Chatila, R., Chetouani, M., Alami, R.: Robots learning how and where to approach people. In: 2016 25th IEEE international symposium on robot and human interactive communication (RO-MAN). pp. 347–353. IEEE (2016) 8

[93] Ravichandar, H.C., Dani, A.P.: Human intention inference using expectation-maximization algorithm with online model learning. IEEE Transactions on Automation Science and Engineering 14(2), 855–868 (2016) 4

[94] Rekleitis, I.M., Dudek, G., Milios, E.E.: Multi-robot exploration of an unknown environment, efficiently reducing the odometry error. In: International joint conference on artificial intelligence. vol. 15, pp. 1340–1345. LAWRENCE ERLBAUM ASSOCIATES LTD (1997) 10

[95] Rios-Martinez, J., Spalanzani, A., Laugier, C.: From proxemics theory to socially-aware navigation: A survey. International Journal of Social Robotics 7(2), 137–153 (2015) 5, 8, 12

[96] Rossi, F., Bandyopadhyay, S., Wolf, M., Pavone, M.: Review of multi-agent algorithms for collective behavior: a structural taxonomy. IFAC-PapersOnLine 51(12), 112–117 (2018) 9

[97] Rudenko, A., Palmieri, L., Herman, M., Kitani, K.M., Gavrila, D.M., Arras, K.O.: Human motion trajectory prediction: A survey. arXiv preprint arXiv:1905.06113 (2019) 5

[98] Ruys, K.I., Aarts, H.: When competition merges people's behavior: Interdependency activates shared action representations. Journal of Experimental Social Psychology 46(6), 1130–1133 (2010) 4

[99] Salmela, M., Nagatsu, M.: How does it really feel to act together? shared emotions and the phenomenology of we-agency. Phenomenology and the Cognitive Sciences 16(3), 449–470 (2017) 4

[100] Satake, S., Kanda, T., Glas, D.F., Imai, M., Ishiguro, H., Hagita, N.: How to approach humans?: strategies for social robots to initiate interaction. In: Proceedings of the 4th ACM/IEEE international conference on Human robot interaction. pp. 109–116. ACM (2009) 8

[101] Sebanz, N., Bekkering, H., Knoblich, G.: Joint action: bodies and minds moving together. Trends in cognitive sciences 10(2), 70–76 (2006) 4

[102] Sebanz, N., Knoblich, G., Prinz, W.: How two share a task: corepresenting stimulus-response mappings. Journal of Experimental Psychology: Human Perception and Performance 31(6), 1234 (2005) 4

[103] Silvagni, M., Tonoli, A., Zenerino, E., Chiaberge, M.: Multipurpose uav for search and rescue operations in mountain avalanche events. Geomatics, Natural Hazards and Risk 8(1), 18–33 (2017) 7

[104] Simon, J.R., Wolf, J.D.: Choice reaction time as a function of angular stimulus-response correspondence and age. Ergonomics 6(1), 99–105 (1963) 4

[105] Singer, S., Akin, D.: A survey of quantitative team performance metrics for human-robot collaboration. In: 41st International Conference on Environmental Systems. p. 5248 (2011) 9

[106] Sreedharan, S., Chakraborti, T., Muise, C., Kambhampati, S.: Expectation-aware planning: A unifying framework for synthesizing and executing self-explaining plans for human-aware planning (2020) 7

[107] Stalnaker, R.: Common ground. Linguistics and philosophy 25(5), 701–721 (2002) 4

[108] Štolba, M., Urbanovská, M., Fišer, D., Komenda, A.: Cost partitioning for multi-agent planning. In: Proceedings of the 11th International Conference on Agents and Artificial Intelligence, ICAART (2019) 6, 11

[109] The, V.N., Jayawardena, C.: A decision making model for optimizing social relationship for side-by-side robotic wheelchairs in active mode. In: IEEE International Conference on Biomedical Robotics and Biomechatronics (BioRob). pp. 735–740 (2016) 8

[110] Thomason, J., Zhang, S., Mooney, R.J., Stone, P.: Learning to interpret natural language commands through human-robot dialog. In: Twenty-Fourth International Joint Conference on Artificial Intelligence (2015) 6

[111] Tomasello, M., Carpenter, M.: Shared intentionality. Developmental science 10(1), 121–125 (2007) 4

[112] Torreño, A., Onaindia, E., Komenda, A., Štolba, M.: Cooperative multi-agent planning: A survey. ACM Computing Surveys (CSUR) 50(6), 1–32 (2017) 9

[113] Tsai, C.C., Kuo, W.J., Hung, D.L., Tzeng, O.J.: Action co-representation is tuned to other humans. Journal of Cognitive Neuroscience 20(11), 2015–2024 (2008) 4

[114] Turnwald, A., Althoff, D., Wollherr, D., Buss, M.: Understanding human avoidance behavior: interaction-aware decision making based on game theory. International Journal of Social Robotics 8(2), 331–351 (2016) 5

[115] Van Ginkel, W., Tindale, R.S., Van Knippenberg, D.: Team reflexivity, development of shared task representations, and the use of distributed information in group decision making. Group Dynamics: Theory, Research, and Practice 13(4), 265 (2009) 4

[116] Van Ginkel, W., Van Knippenberg, D.: Group information elaboration and group decision making: The role of shared task representations. Organizational behavior and human decision processes 105(1), 82–97 (2008) 4

[117] Van Welie, M., Van der Veer, G.C., Eliëns, A.: An ontology for task world models. In: Design, Specification and Verification of Interactive Systems' 98, pp. 57–70. Springer (1998) 5

[118] Vasquez, D., Stein, P., Rios-Martinez, J., Escobedo, A., Spalanzani, A., Laugier, C.: Human aware navigation for assistive robotics. In: Experimental Robotics. pp. 449–462. Springer (2013) 8

[119] Vemula, A., Muelling, K., Oh, J.: Social attention: Modeling attention in human crowds. In: 2018 IEEE International Conference on Robotics and Automation (ICRA). pp. 1–7. IEEE (2018) 5

[120] Wei, C., Hindriks, K.V., Jonker, C.M.: The role of communication in coordination protocols for cooperative robot teams. In: ICAART (2). pp. 28–39 (2014) 9, 10

[121] Williams, A., Sebastian, B., Ben-Tzvi, P.: Review and analysis of search, extraction, evacuation, and medical field treatment robots. Journal of Intelligent & Robotic Systems pp. 1–18 (2019) 7

[122] Wykowska, A., Chellali, R., Al-Amin, M.M., Müller, H.J.: Implications of robot actions for human perception. how do we represent actions of the observed robots? International Journal of Social Robotics 6(3), 357–366 (2014) 6, 11

[123] Yuan, F., Hanheide, M., Sagerer, G., et al.: Spatial context-aware person-following for a domestic robot. International Workshop on Cognition for Technical Systems (2008) 8

[124] Zanlungo, F., Ikeda, T., Kanda, T.: Social force model with explicit collision prediction. EPL (Europhysics Letters) 93(6), 68005 (2011) 5

[125] Zhang, H., Luo, H., Wang, Z., Liu, Y., Liu, Y.: Multi-robot cooperative task allocation with definite path-conflict-free handling. IEEE Access 7, 138495–138511 (2019) 10